



# COPC Mapping Refinement



CPG/EPA Meeting

April 27, 2016

# Outline

- Refinements and Other Ideas Examined
- Revised Results
- Model ICs and Remedial Alternatives
- Subsurface
- Proposed Next Steps

# Implemented Refinements

- “River straightening” approach tweaked to better follow the channel centerline
- Groups into which the river is broken aggregated into smaller number
- Variogram used for simulations refined, and now includes a nugget
- Additional simulation QC metrics added

# Other Ideas Examined

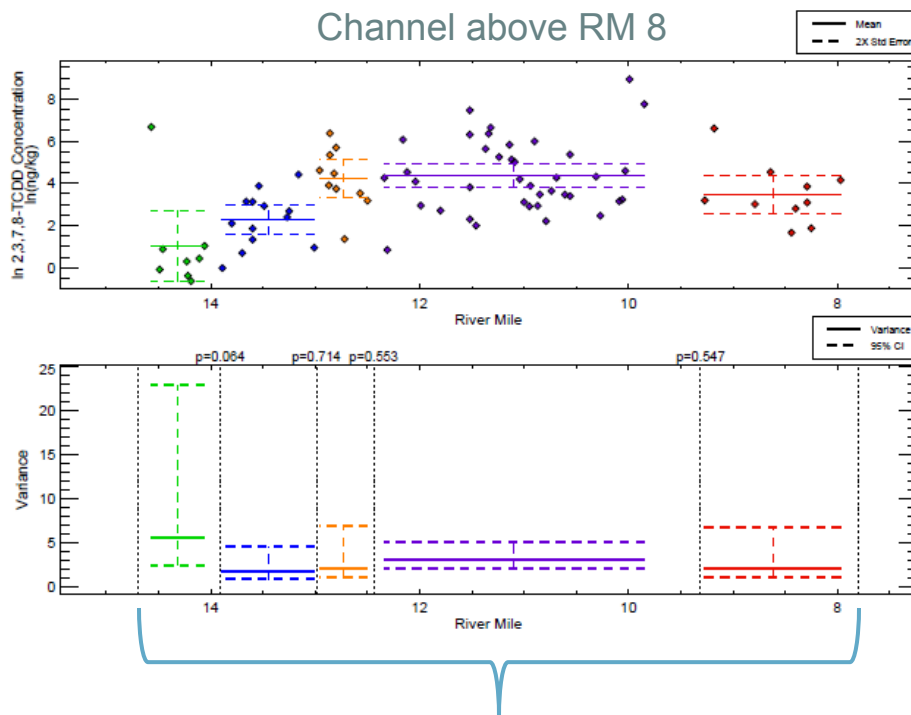
- Check whether concentration correlates strongly with other parameters such that a predictive statistical model can be used to inform interpolation
- Check if historical data can be used to check structure of applied variogram
- Use of data to support anisotropy ratio
- Alternative transformation aimed at achieving normally distributed data (“normal scores transform”)
- Established model initial conditions and remedial alternatives from mapping results
- Mapping of subsurface concentration

# Aggregate Groups

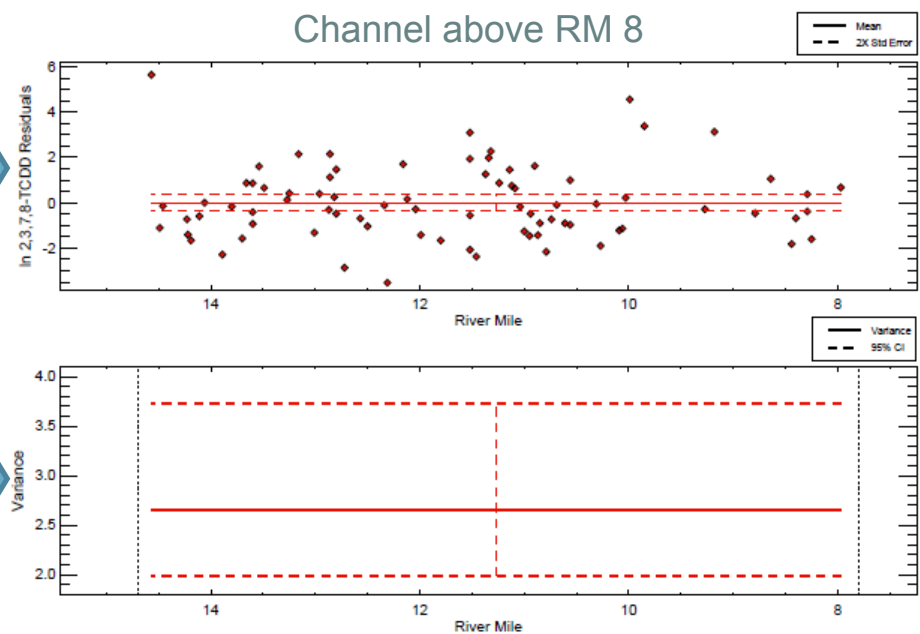
- Motivation
  - Fewer groups increases data count in each group
  - Fewer groups means fewer sharp concentration breaks that are unavoidable at group boundaries
- Approach
  - Combine groups with similar concentration variance
  - Conduct simulations on residual concentration (value – mean)
    - Allows combining of groups with similar variance and differing means (means are added back in after simulation)
  - Reduced 26 groups to 9 groups
    - Sub-groups of differing means (total of 23 “data groups”)

# Example of Data Group Pooling

## Data Groups



## Simulation Groups

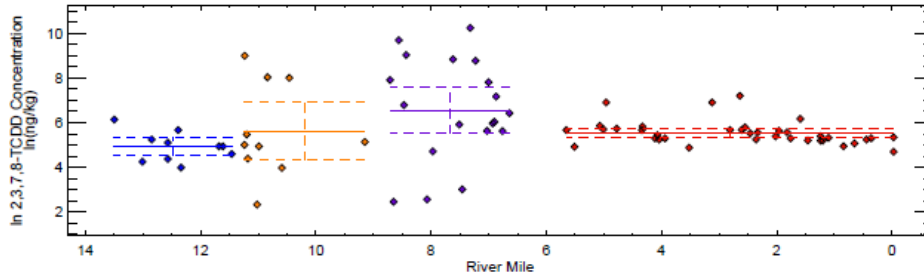


Similar variance, so pool these data groups in the conditional simulations (as mean-removed residual concentrations)

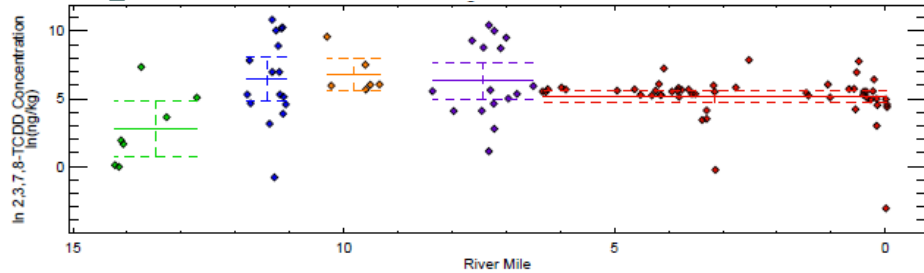
# Summary of Revised Groups

## Data Groups

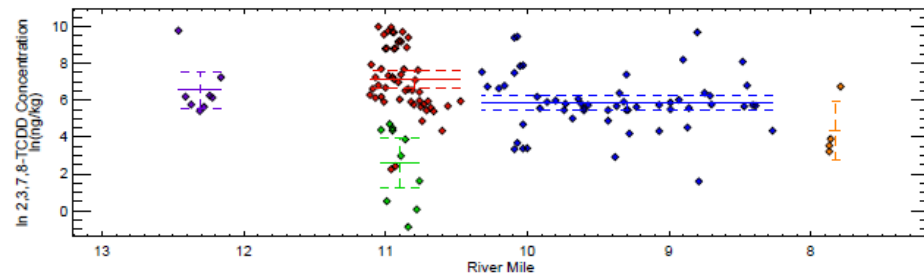
Left shoal



Right shoal

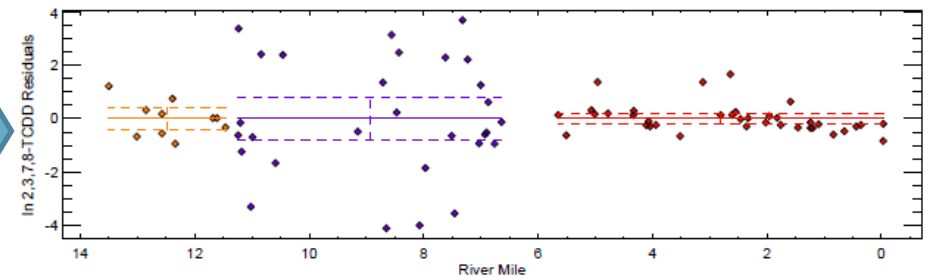


Silt

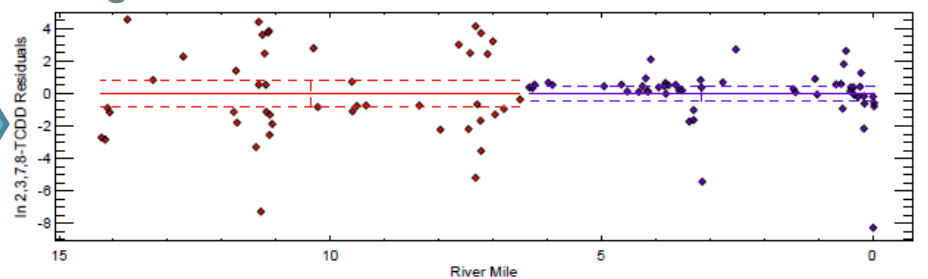


## Simulation Groups

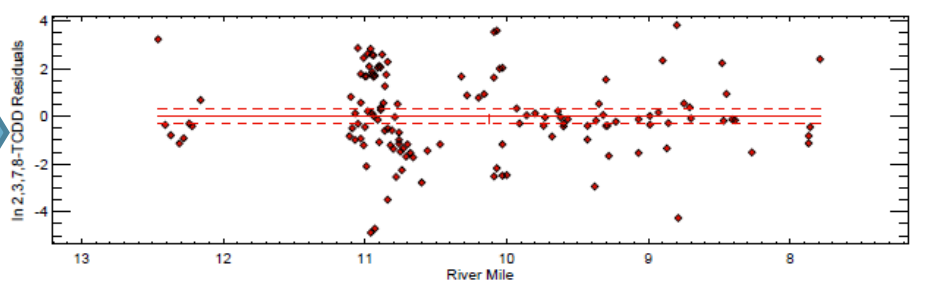
Left shoal



Right shoal



Silt

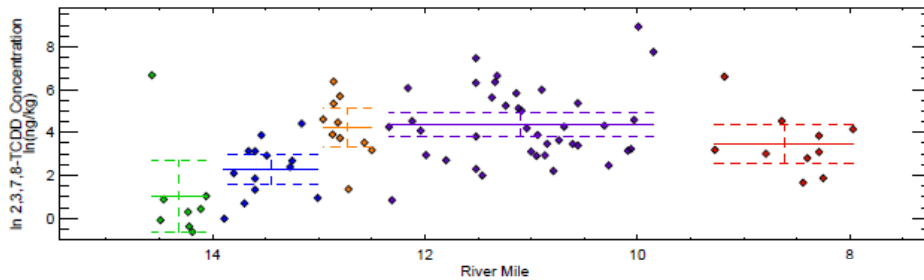




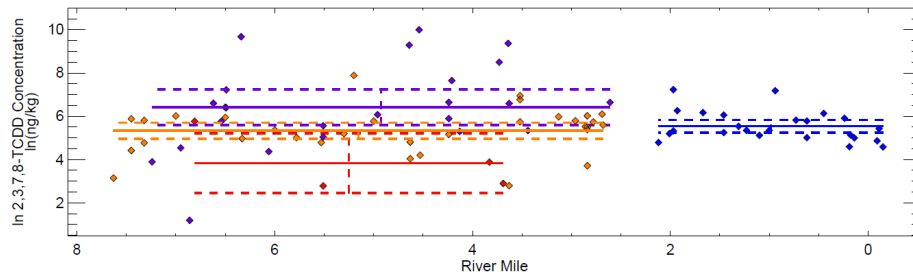
# Summary of Revised Groups

## Data Groups

Channel above RM 8

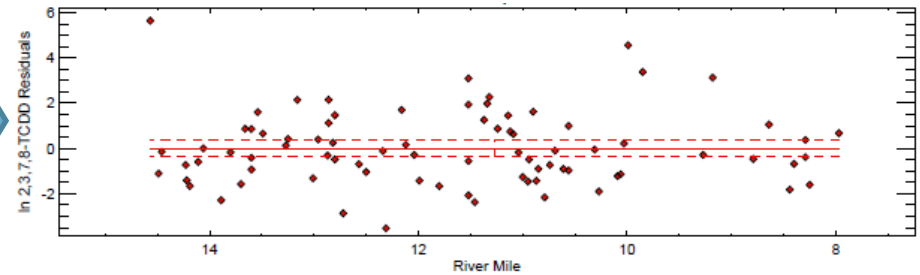


Channel below RM 8

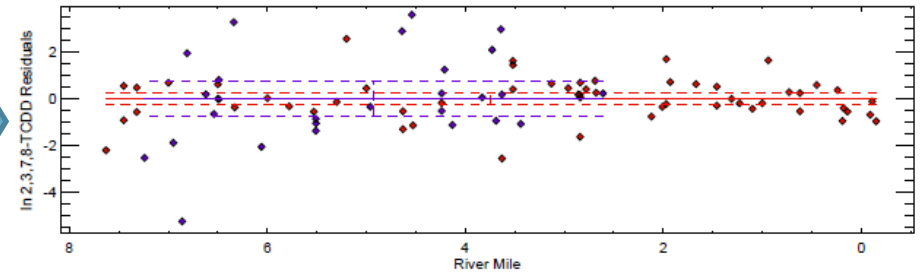


## Simulation Groups

Channel above RM 8



Channel below RM 8

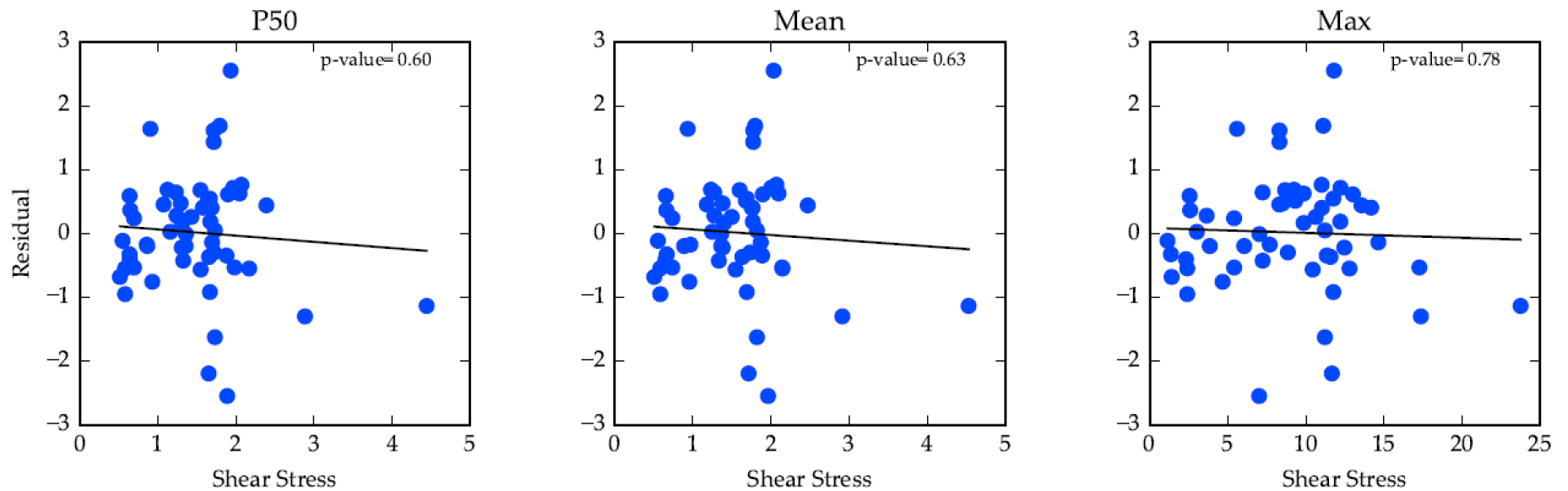




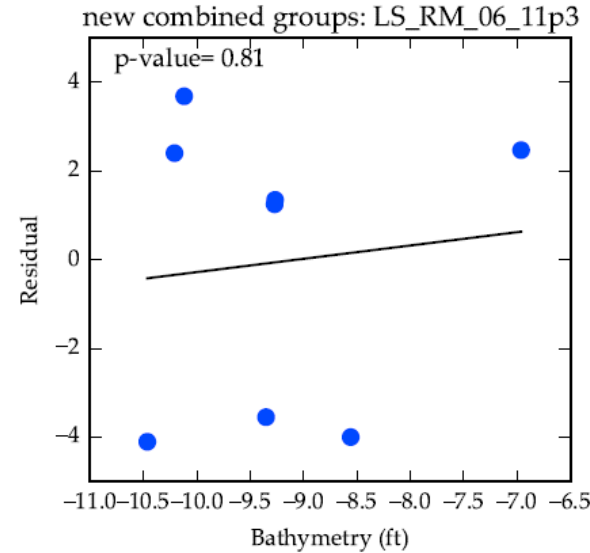
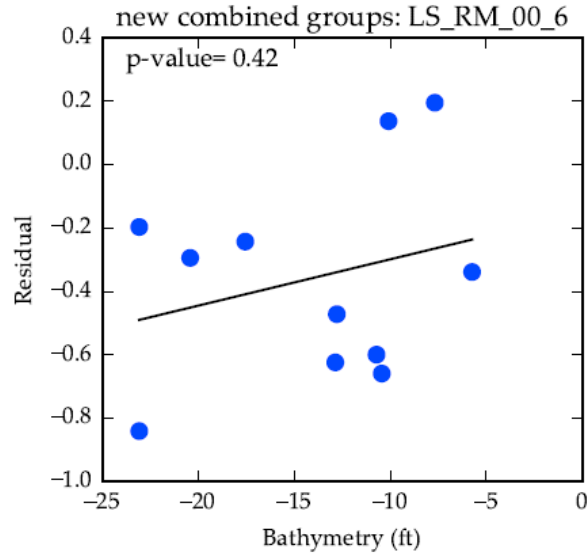
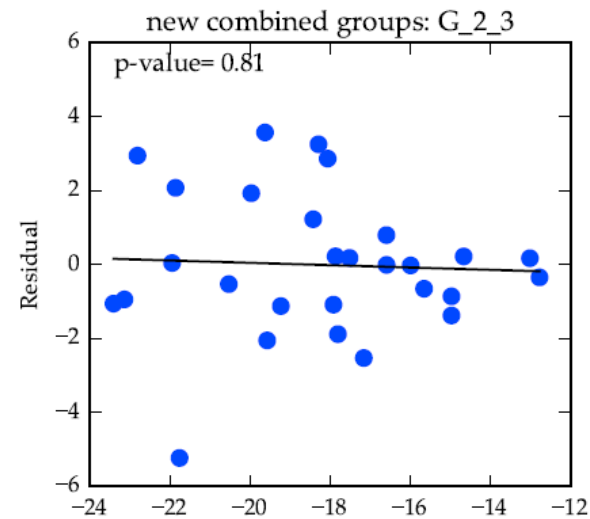
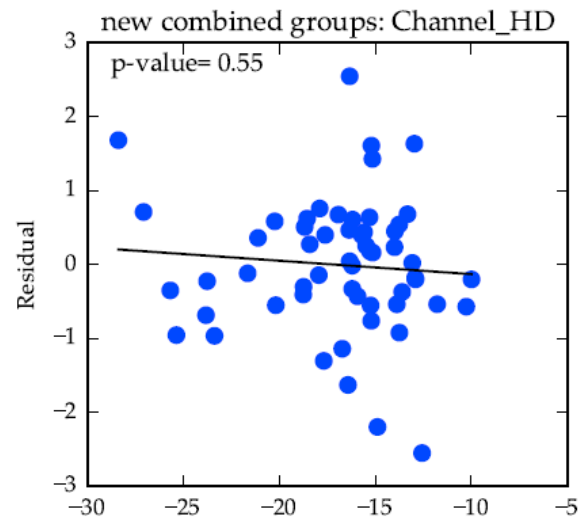
# Correlations to In(C) Residuals in Pooled Simulation Groups

- Shear stress (high resolution hydrodynamic model):
  - no significant correlation at the 5% confidence level for P50, mean, and max shear stress
- Bathymetry:
  - 2011 bathymetry (Irene): no significant correlation at the 5% confidence level
  - Looking at 1949, 1966, 1995 – 1999, 2001, 2004, 2007, 2008, 2010, 2011, and 2012 bathymetry: 9 out of 86 (~10%) of the data sets had significant correlation
    - Suspect that effect is already captured in part by the spatially variable means
- Side scan sonar sediment type:
  - only 6 out of 46 tested significant and
  - of the 6 only 2 had more than 5 data points

# Shear Stress Example

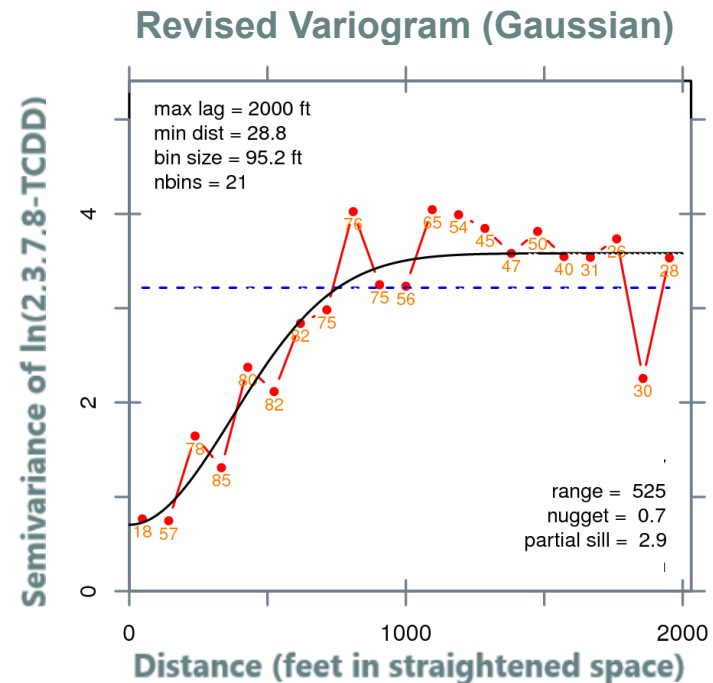
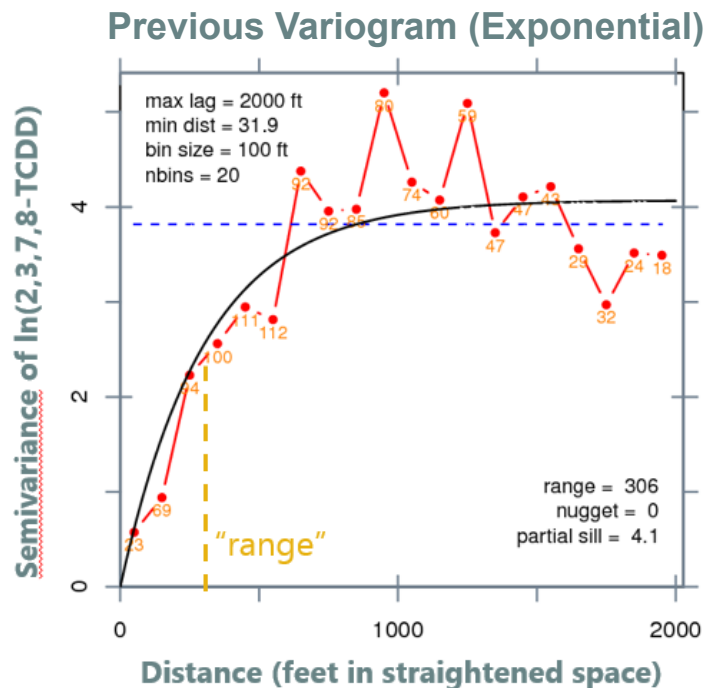


# 2011 Bathymetry Example



# Revision of the Variogram (Along-flow)

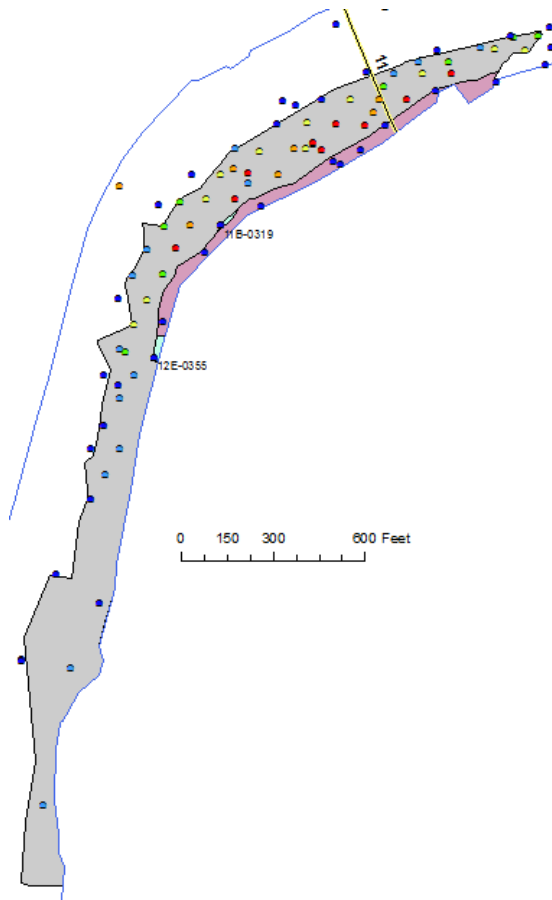
- As before, variogram shape is derived from the RM 10.9 data, and scaled to other groups using local variance
- The fit is now Gaussian and it includes a nugget
- Minor changes to the underlying dataset (next slide)



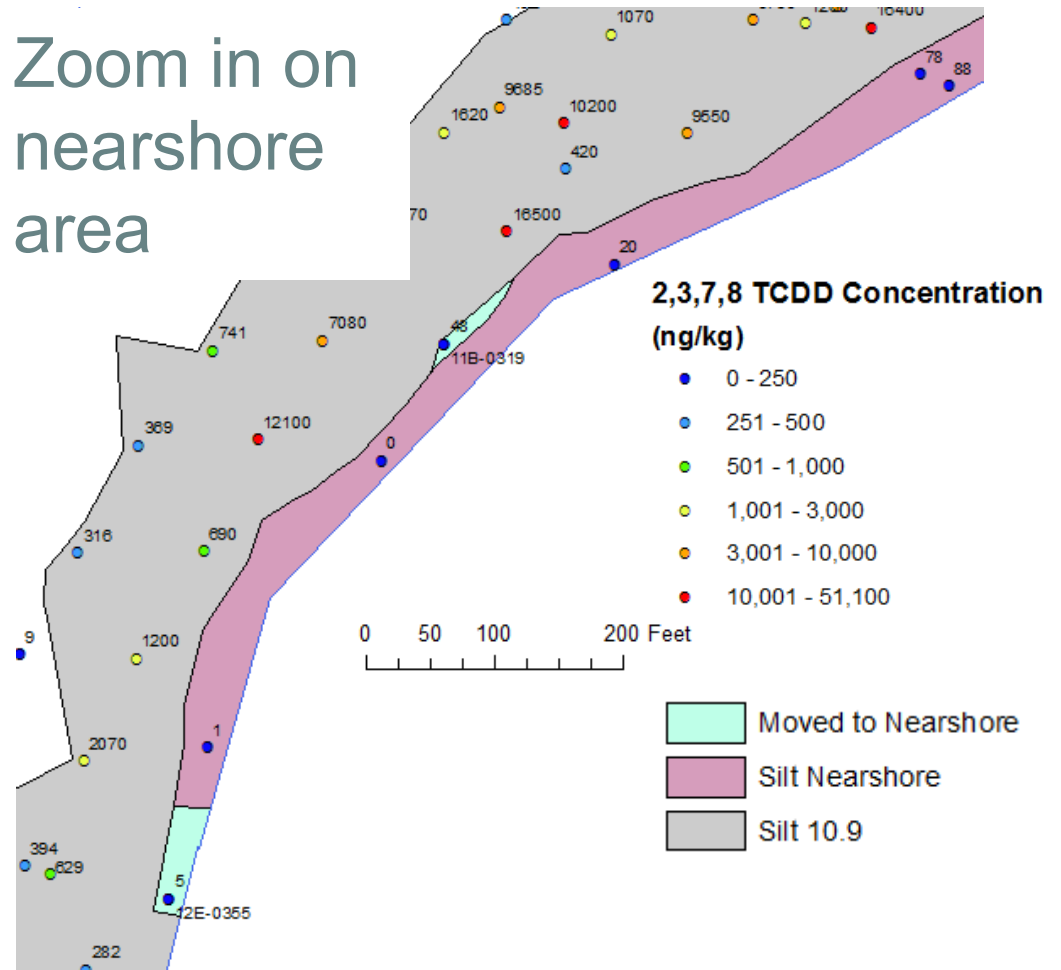
# Revision of the Variogram

## Edit to RM 10.9 Silt Deposit Delineation

RM 10.9 Silt Deposit



Zoom in on nearshore area



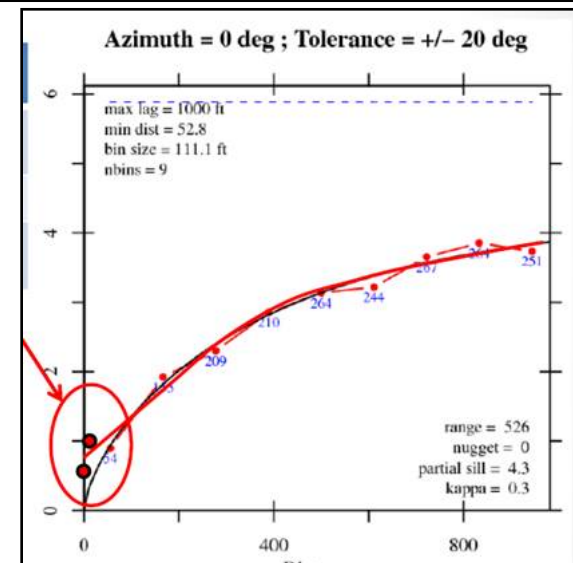
# Revision of the Variogram - Nugget

- Variogram nugget is consistent with prior EPA analysis
  - CPG revised variogram has a nugget of 0.7, which is about 20% of sill
  - EPA analysis suggested a nugget of 0.57 to 1.01, and 20 to 30% of sill
- CPG did not attempt to recreate EPA analysis due to dataset and data treatment differences
  - EPA used data from whole river for 2007-2012, included data segmented finer than 6", and did not stratify by groups or include directionality

Excerpts from March 2015 EPA Presentation

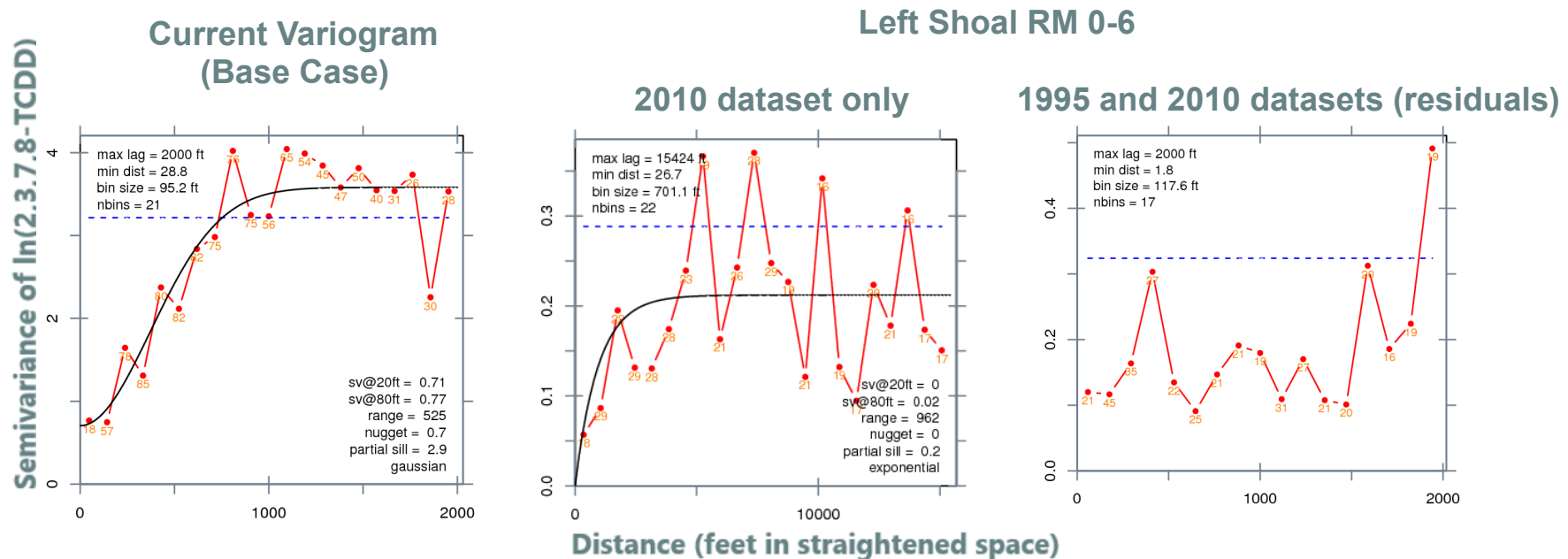
Reanalysis of Close Proximity Samples			
Metric	Number of Pairs	Average <sup>1</sup> Distance (ft)	Semi-variance Log(ng/kg) <sup>2</sup>
Log(TCDD)	25	0	0.57
Log(TCDD)	46	9.2 (2.9-20)	1.01

• Nugget effect on the order of 20% to 30% of the sill



# Variograms Using Other Datasets – Left Shoal

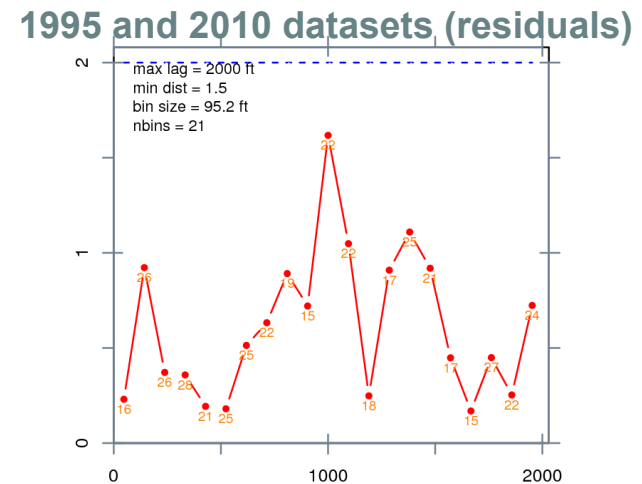
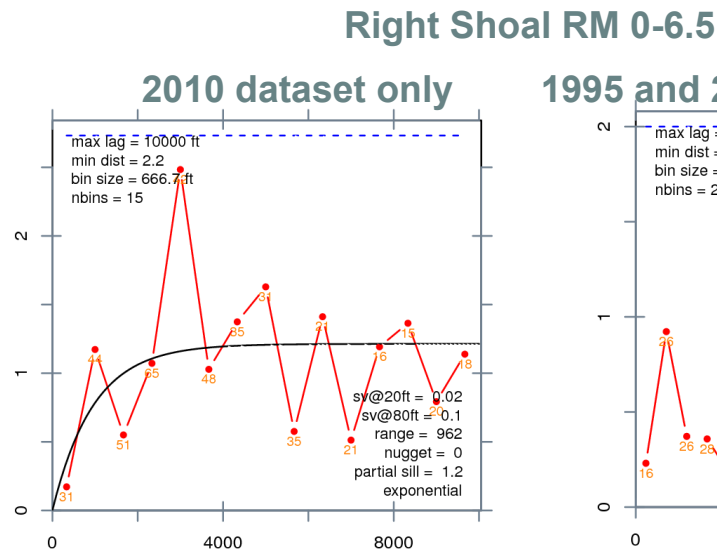
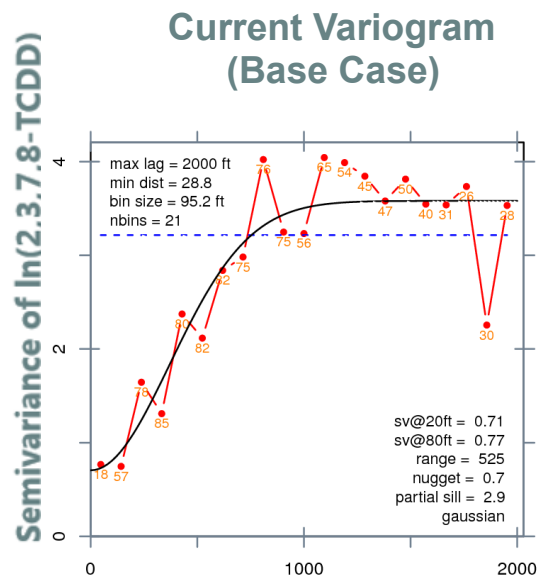
- Left shoal variogram based on only the 2010 dataset suggests long-scale spatial correlation, but lacks closely spaced data
- Adding the 1995 dataset on a residuals basis yields more data pairs at small separation distances. However, the combined dataset lacks spatial correlation, which may be an artifact of combining datasets from two different time periods on a residuals basis





# Variograms Using Other Datasets – Right Shoal

- Right shoal variogram based on only the 2010 dataset suggests long-scale spatial correlation, but lacks closely spaced data
- Adding the 1995 dataset on a residuals basis yields more data pairs at small separation distances. However, the combined dataset lacks spatial correlation, which may be an artifact of combining datasets from two different time periods on a residuals basis



Distance (feet in straightened space)

# Cross-Flow Variogram

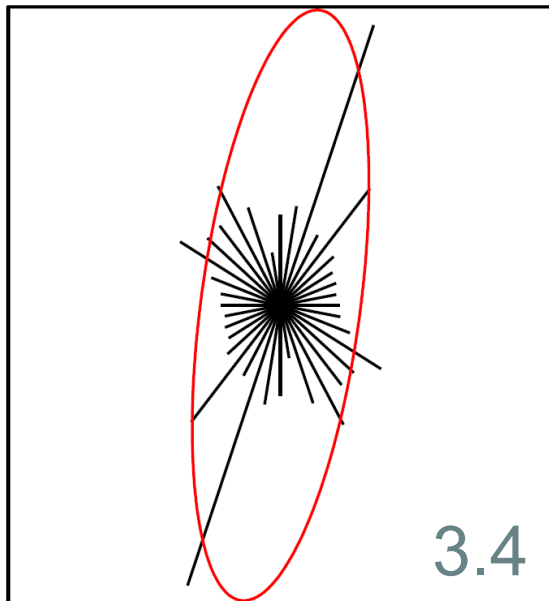
- As before, the along-flow variogram is scaled using an anisotropy ratio of 5
  - Ratio used by EPA in 2015 COPC mapping critique
- Improved ratio estimates were not identified
  - Sparse data cause estimates of anisotropy to be highly variable and sensitive to data treatment
  - Unclear impact of differences between variogram data analysis and its application within simulations
    - Along-flow variogram from data is based on a +/- 20 degree search band, and so influenced by cross-flow anisotropy
    - Simulation applies the input variograms precisely in along-flow and cross-flow directions

# Anisotropy Ratio – Rose Diagrams

- Ratio of 5 is within the range of rose diagrams below

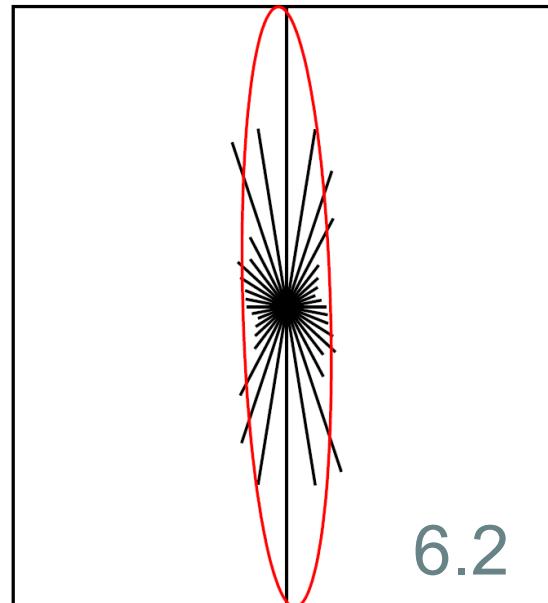
RM 10.9 Silt  
Deposit

Tolerance: 30

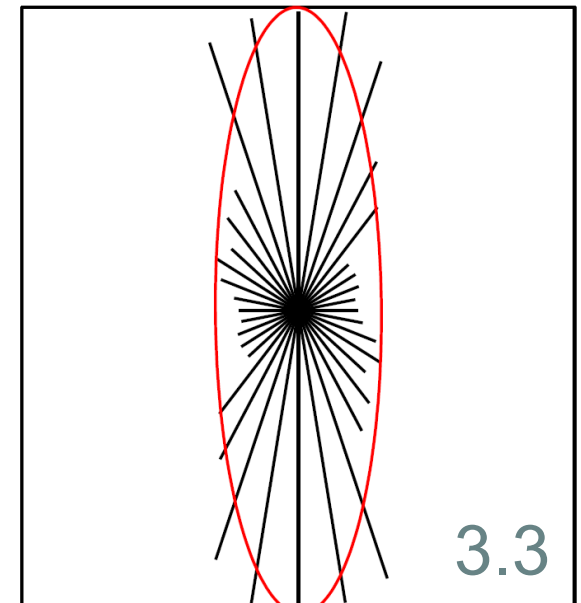


All Data RM 10.4-11.5

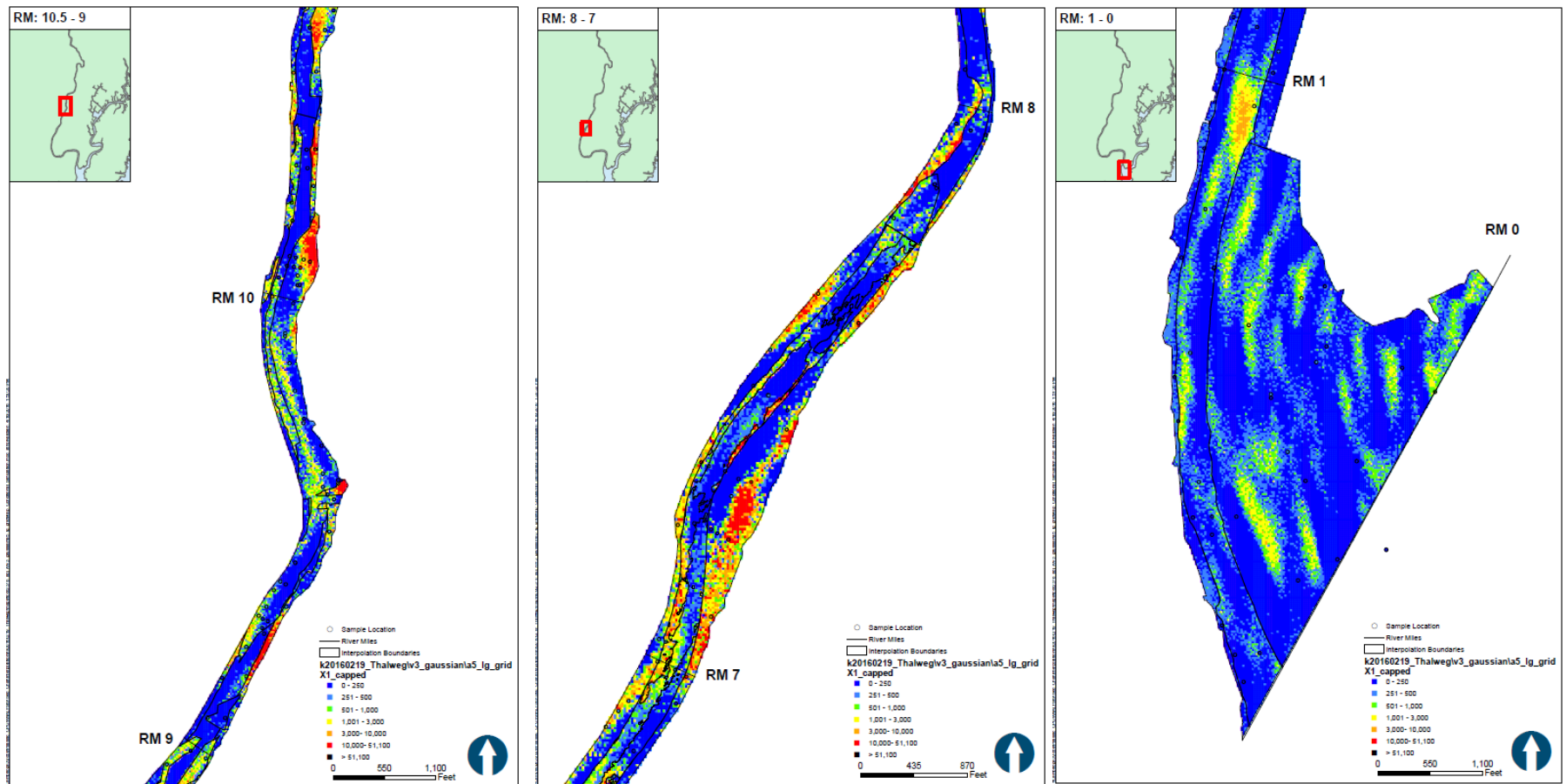
Tolerance: 20



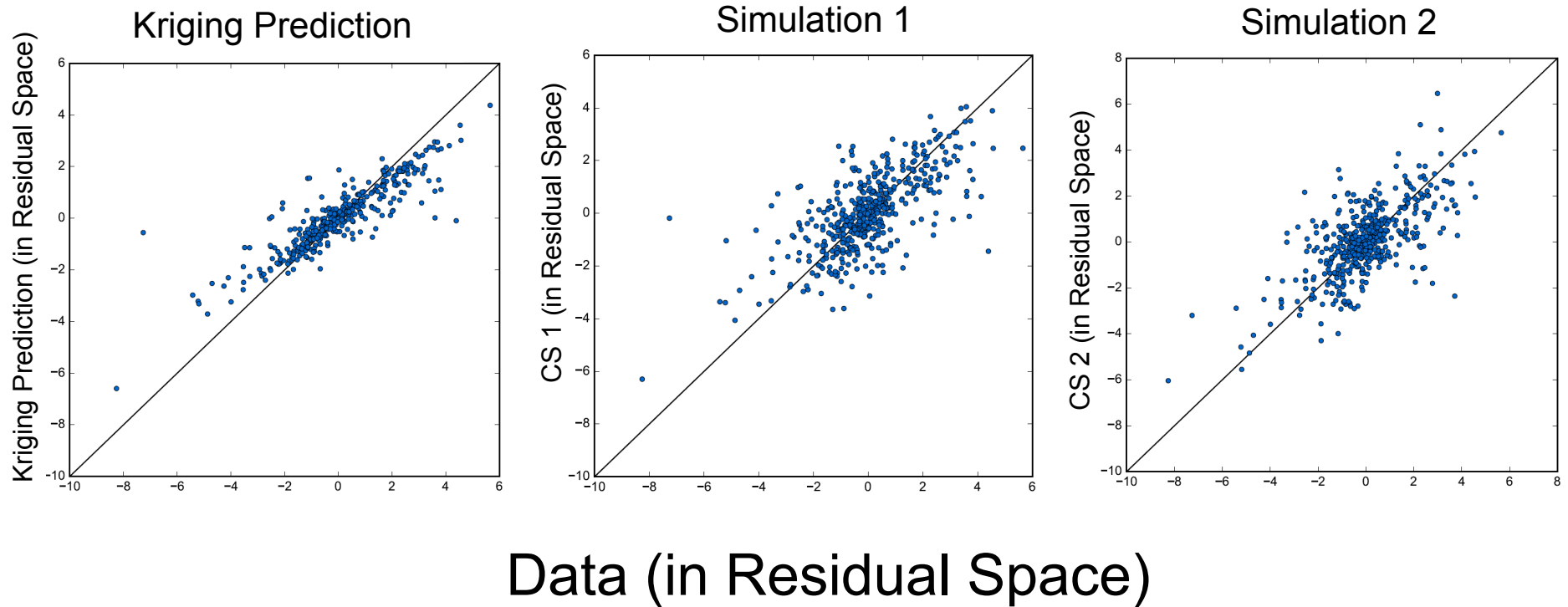
Tolerance: 30



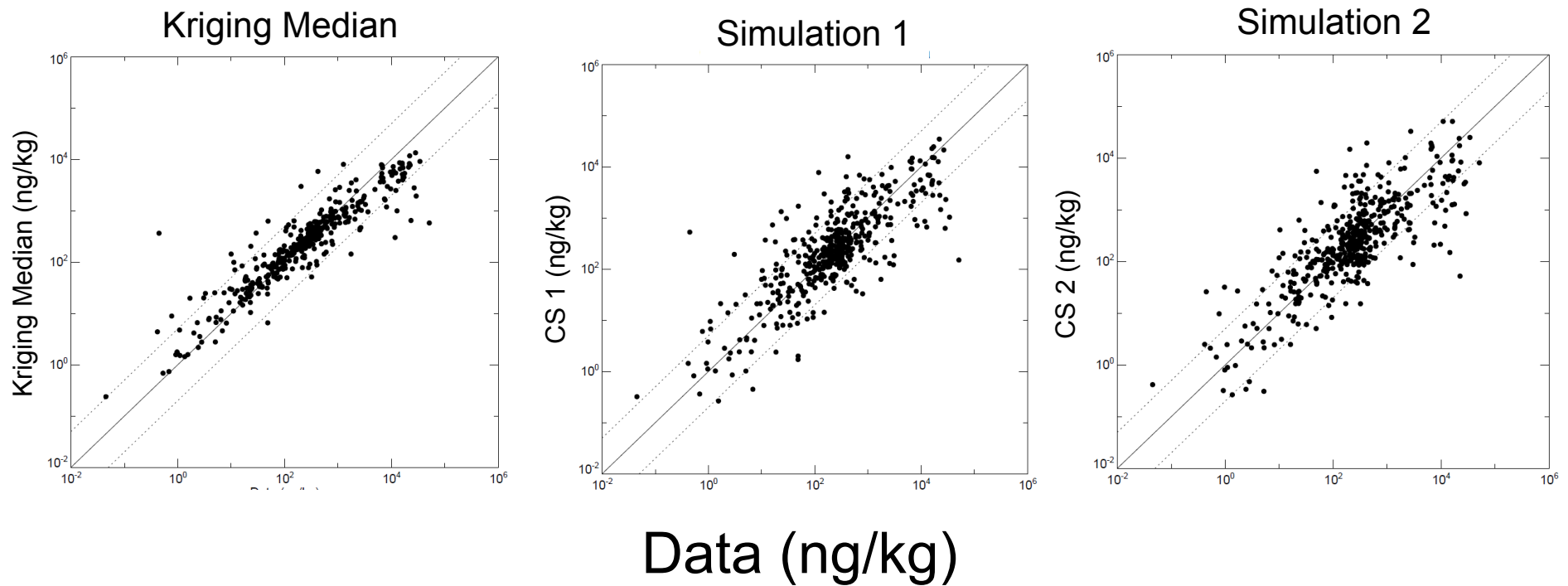
# QC of Results – Sample Maps



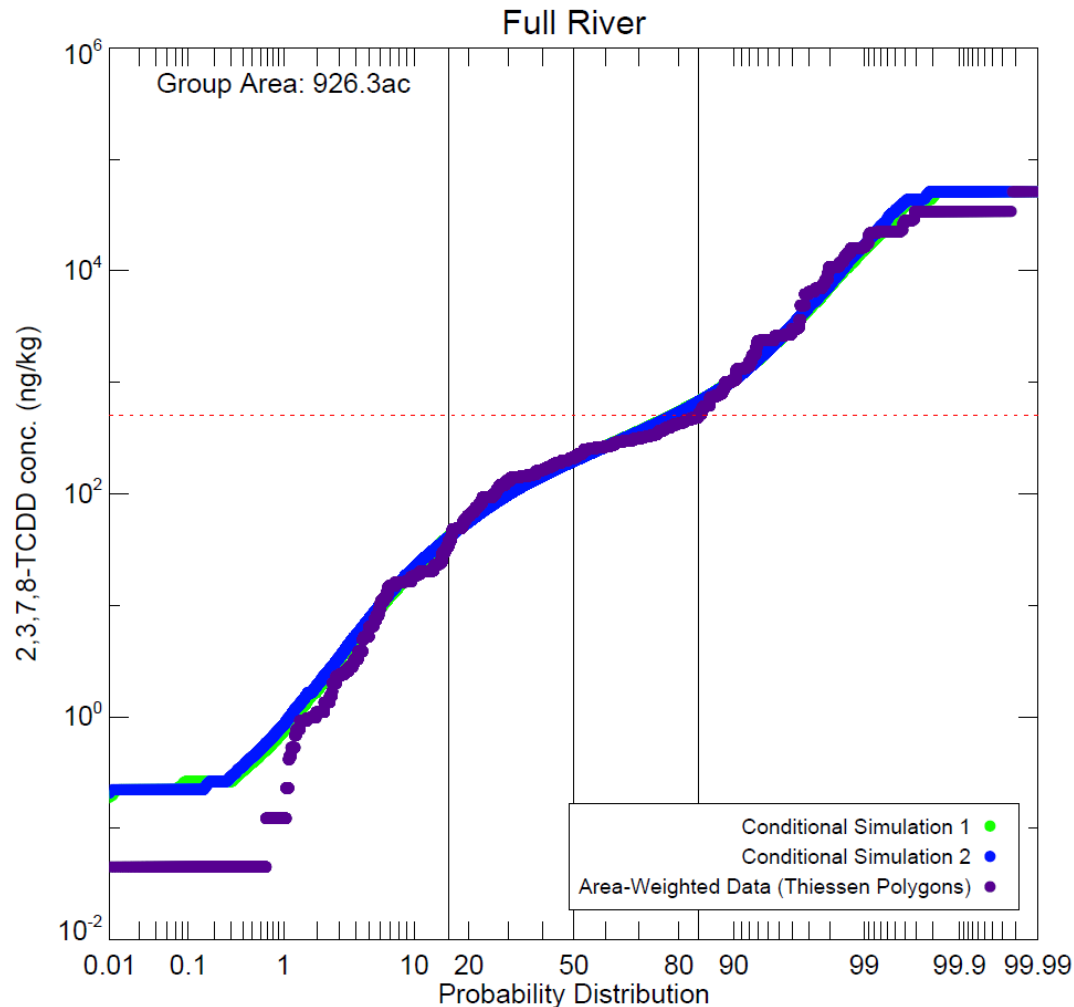
# QC of Residuals (in Natural Log Space)



# QC of Back Transformed Results



# QC of Results – Reproducing Distributions

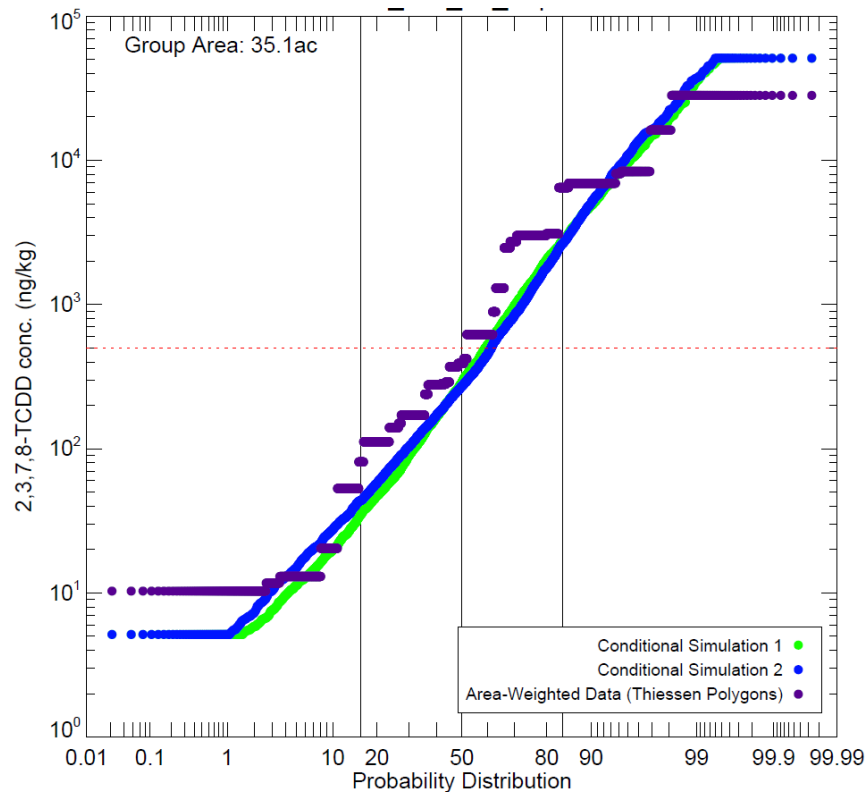


- Concentration caps
  - Max at 51,100 ng/kg or 2x highest measured value in group (whichever is smaller)
  - Min at  $\frac{1}{2}$  lowest measured value in group

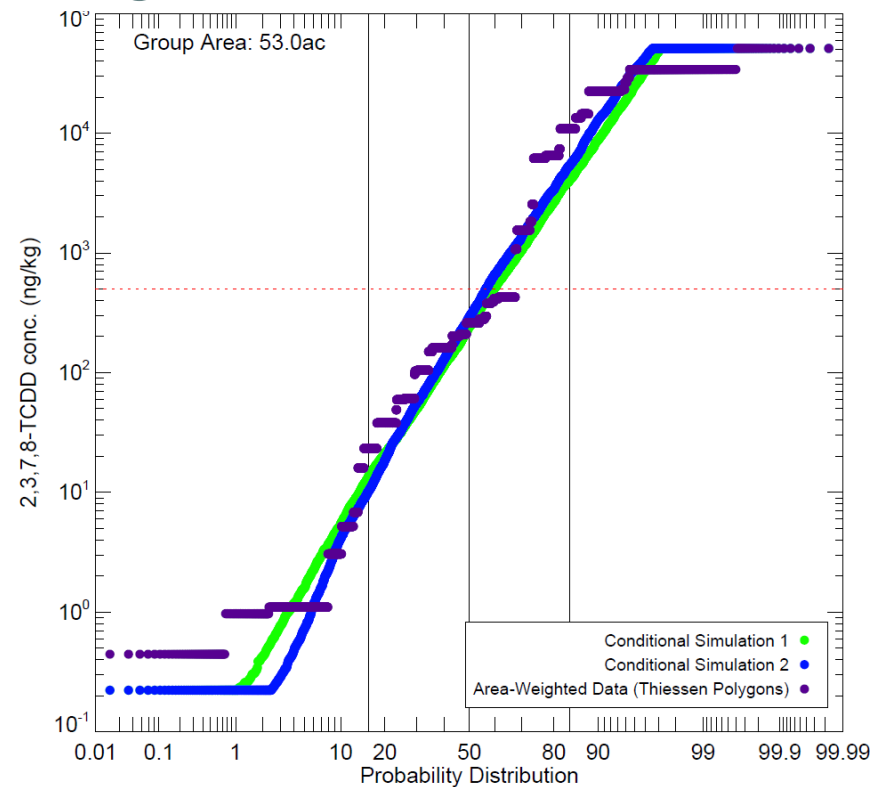


# QC of Results – Reproducing Distributions

## Left Shoal RM 6-11.3

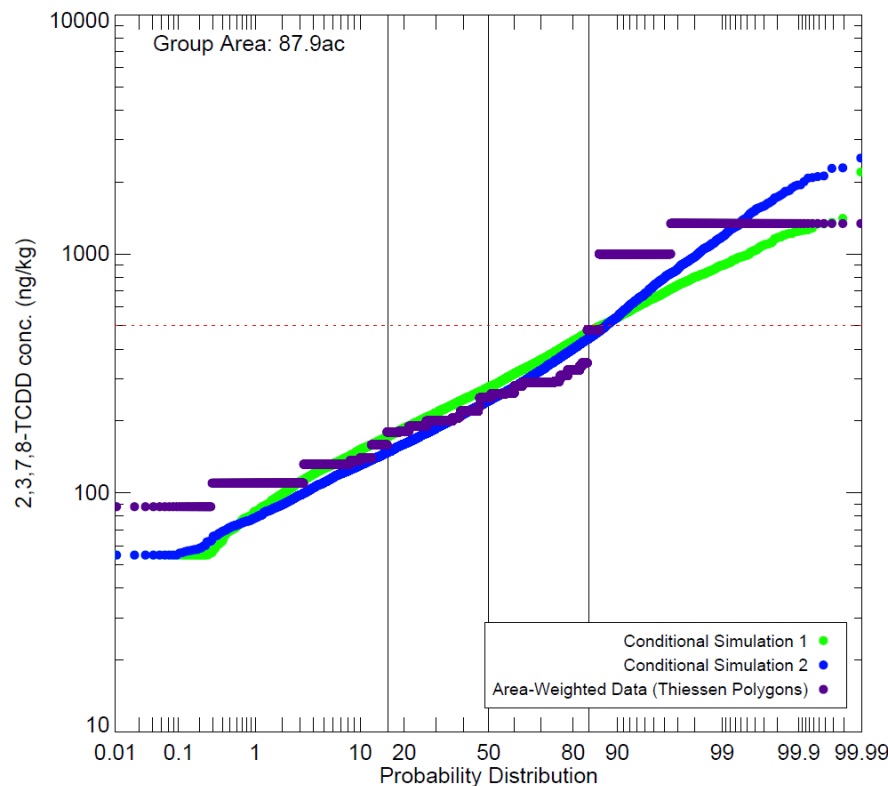


## Right Shoal Above RM 6.5

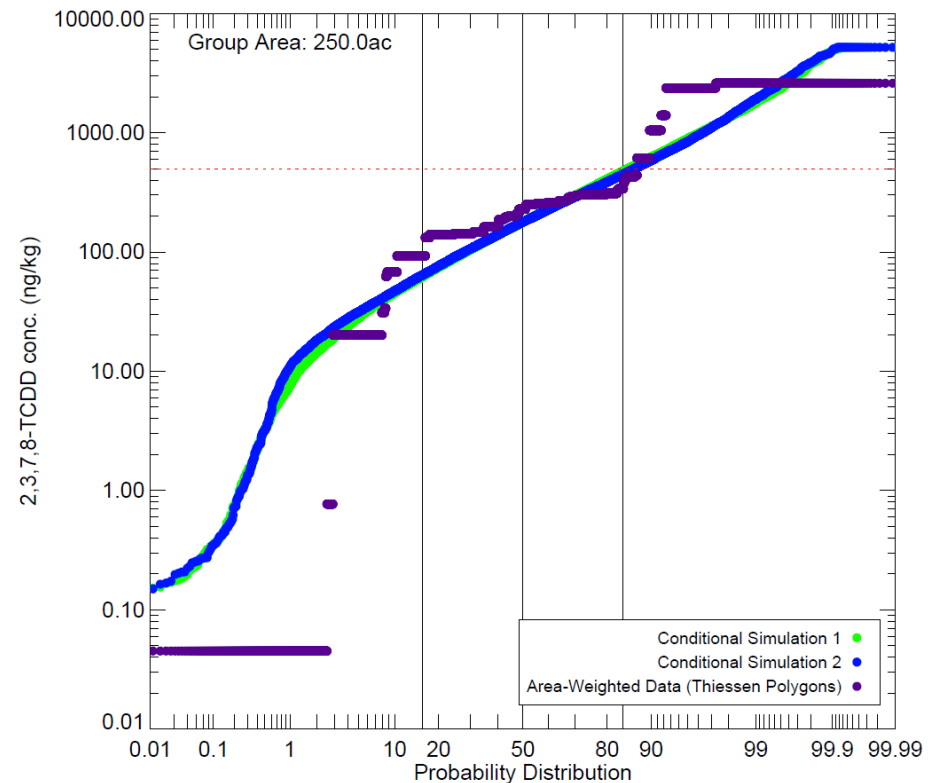


# QC of Results – Reproducing Distributions

## Left Shoal RM 0-6



## Right Shoal Below RM 6.5

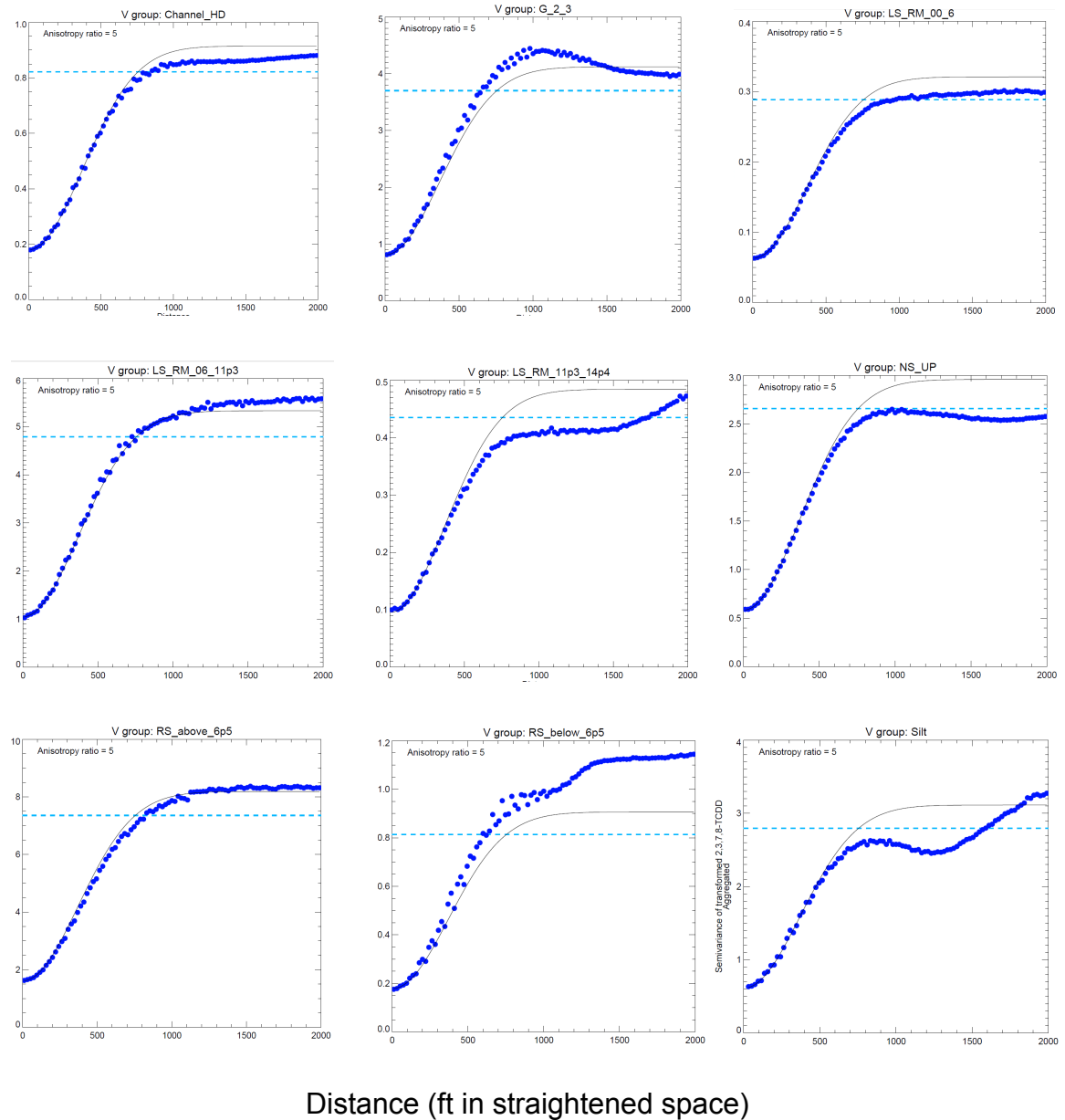


Variance adjusted in Right Shoal RM 0-6.5  
to exclude the two lowest data values

# QC of Results Recovering Variogram

- Generally good recovery at small separation distances

Semivariance of transformed 2,3,7,8-TCDD

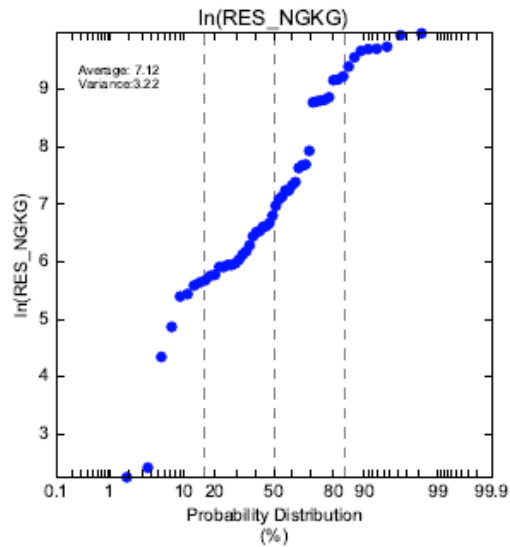


# Alternate Data Transformation (Normal Scores)

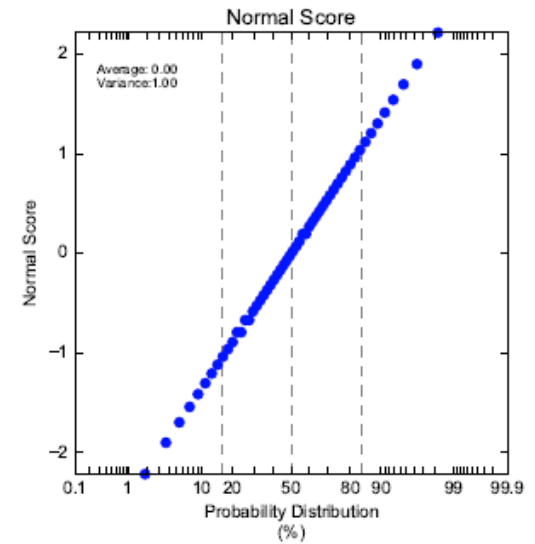
- Conditional simulations assume that input data are normally distributed
- Some non-normality remains after log transforming data
  - Departure from normality varies by simulation group
- A normal score transform based mapping was tested as an alternative
  - The transform forces the data to be normal
- Not selected as base case, mainly due to variogram concerns

# Normal Scores Variogram Comparison

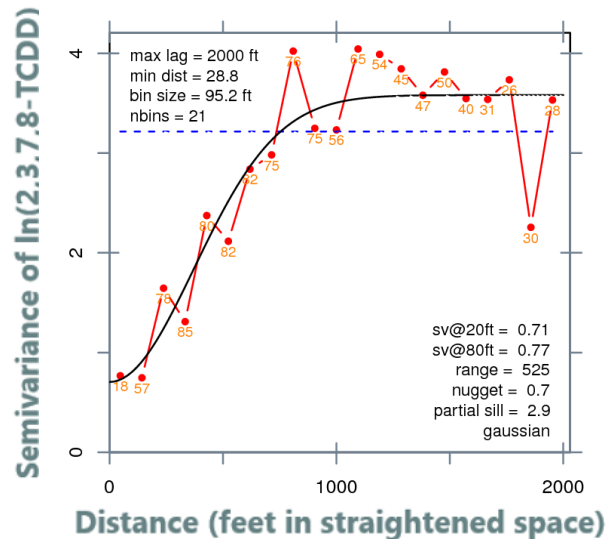
Natural Log Transform  
(Base Case)



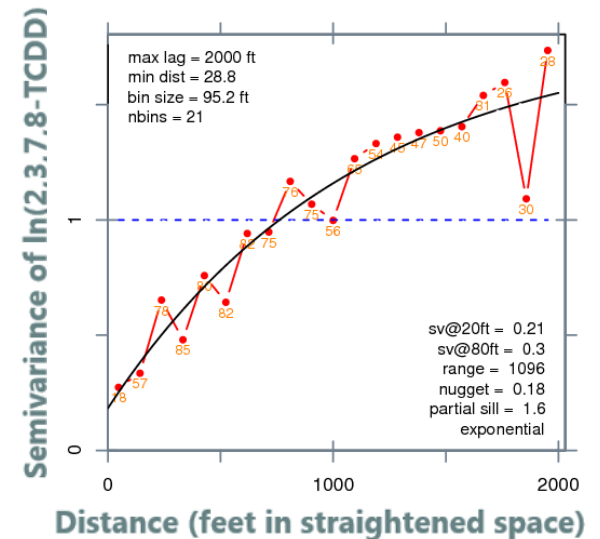
Normal Scores Transform  
(Alternate)



Gaussian Variogram  
(Base Case)



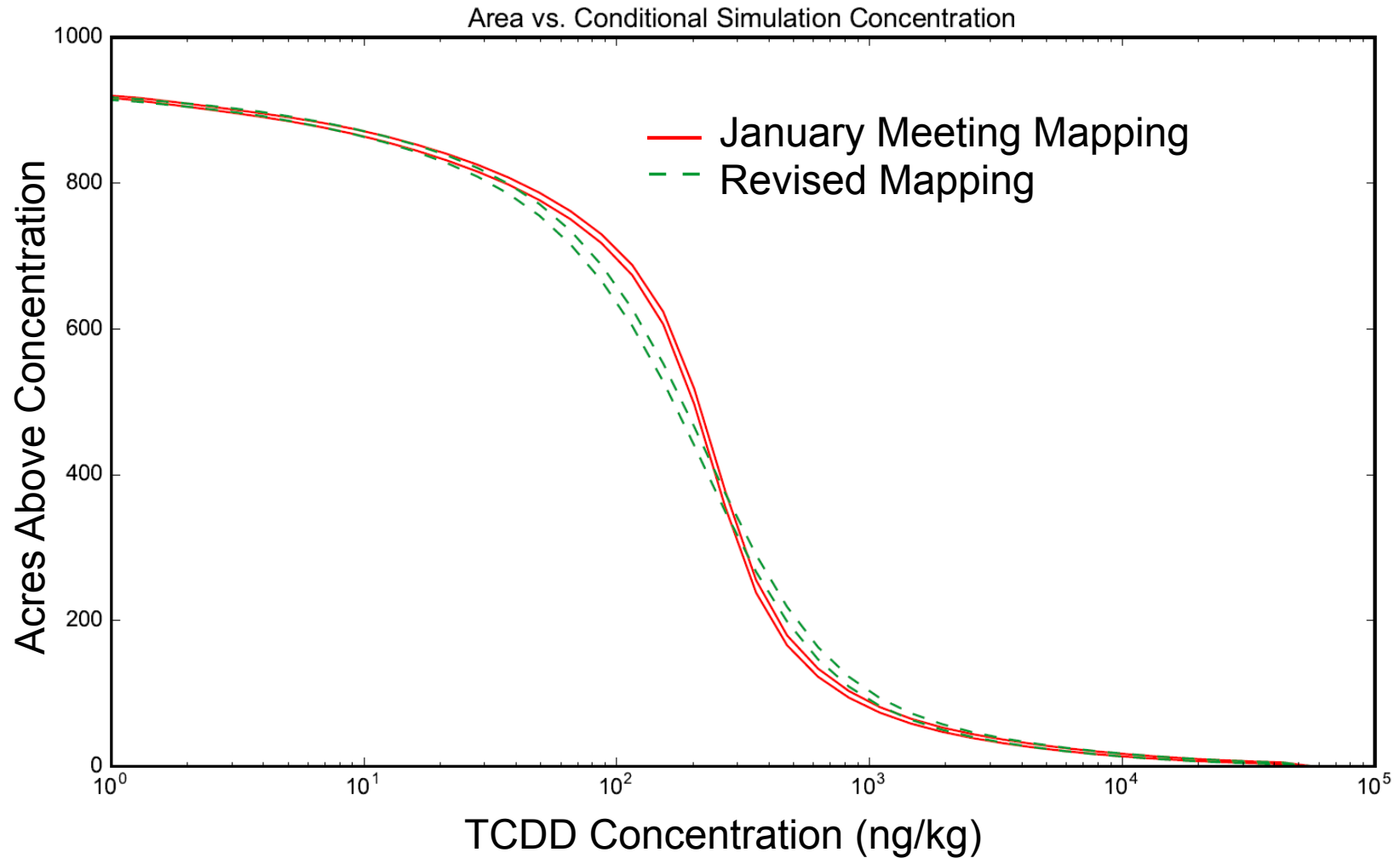
Normal Scores Variogram  
(Alternate)



# Concerns with Using Normal Scores for Mapping

- Shape of the normal scores variogram
  - Approaches the sill much slower than variograms derived in log space; causes large range
- Poor variogram recovery in test simulations
  - Far-scale variance under-predicted in most groups
- Requires applying the same variogram to many groups that each have a different normal score back-transform

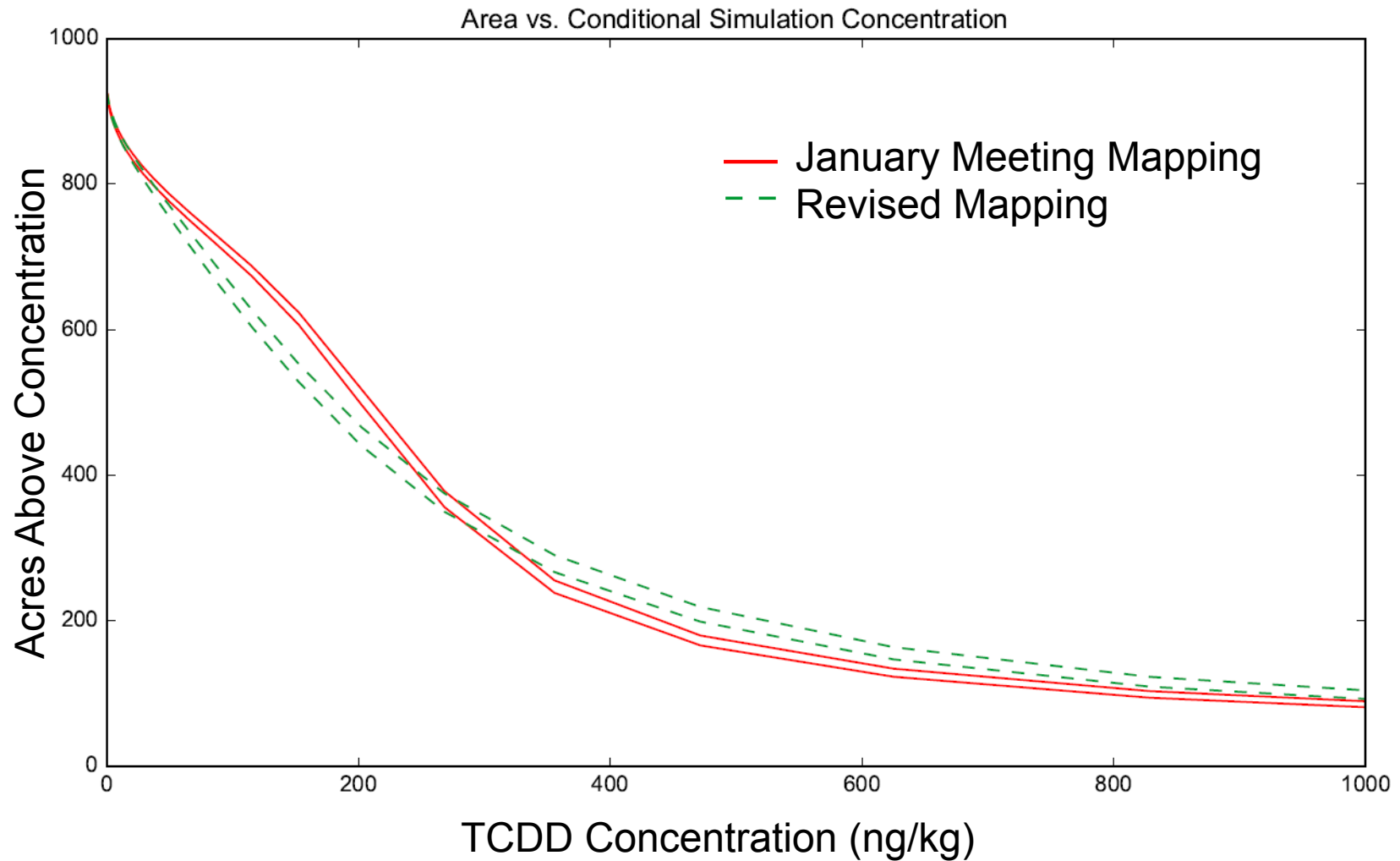
# Revised Mapping Results



*The lines represent +/- 2 SD Envelope*



# Revised Mapping Results (Zoomed In)



*The lines represent  $\pm 2$  SD Envelope*

# Remedial Benefit Evaluation for FS

- FS Needs
    - CFT model initial conditions
    - Realistic concentration change on the scale of the CFT model due to smaller decision unit scale remediation to a given RAL
  - Two basic options for FS discussed in last meeting
    - Use simulation-specific map and footprint
    - Use a single map and footprint
  - Simulation-specific option favored given future design data, but requires picking representative map(s)
    - Must limit to a few given CFT model run-time/effort
  - FS should be based on a fair representation of the conditional simulations
-

# Approach to Craft Remedial Options for FS

- CPG proposes revised approach that
  - Combines all conditional simulations into a single map
  - Preserves remedial benefit for FS evaluations
- Use the average concentration of the simulations for the CFT model initial conditions
  - Avoids having to select a subset of simulations
  - However, average is too smooth to delineate remedial targets
- Impose average remedial benefit from all simulations
  - For each conditional simulation, calculate remedial benefit and acreage for small scale decision unit (e.g., 80 feet)
  - For each decision unit, aggregate results across all simulations
  - Aggregate decision unit averages onto larger CFT model grid cells
  - Use these to adjust CFT model concentrations during FS simulation

## Example Calculation at the DU-scale

Simulation Number	Predicted Concentration (as multiples of a given RAL)	Dredge?	Mass Dredged (per unit volume)	Post Remedy Conc
1	$\frac{1}{4}$	No	0	$\frac{1}{4}$
2	$\frac{1}{2}$	No	0	$\frac{1}{2}$
3	2	Yes	2	0

Average over the simulations for a single DU:

- DU initial conc:  $(\frac{1}{4} + \frac{1}{2} + 2) / 3 = 0.92$
- Mass dredged:  $(0 + 0 + 2) / 3 = 0.67$
- Post-remedy conc:  $(\frac{1}{4} + \frac{1}{2} + 0) / 3 = 0.25$
- Fractional area dredged:  $(0 + 0 + 1) / 3 = 0.33$

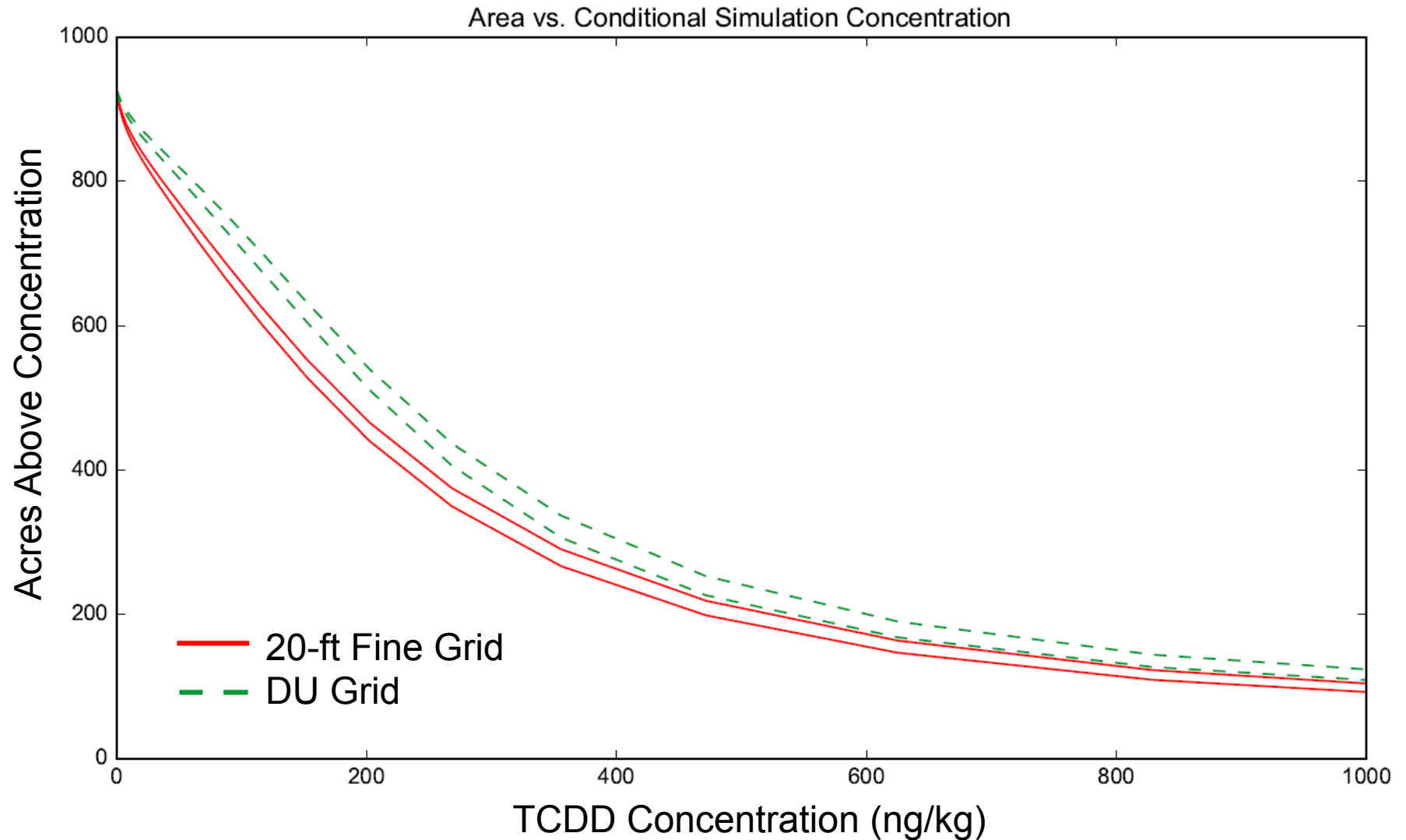
# Example Upscaling of DU Statistics to CFT Model Grid

	<u>Initial Conc</u> DU: Mean of Simulations Model Grid: Area-Weighted Mean of DUs	<u>Post Remedy Conc</u> DU: Mean of Simulations Model Grid: Area-Weighted Mean of DUs	<u>Fractional Area Remediated</u> DU: Mean of Simulations Model Grid: Area-Weighted Mean of DUs
DU1	0.92	0.25	0.33
DU2	1.30	0.17	0.67
DU3	1.90	0.27	0.67
<b>Model Grid</b>	<b>1.37</b>	<b>0.23</b>	<b>0.56</b>

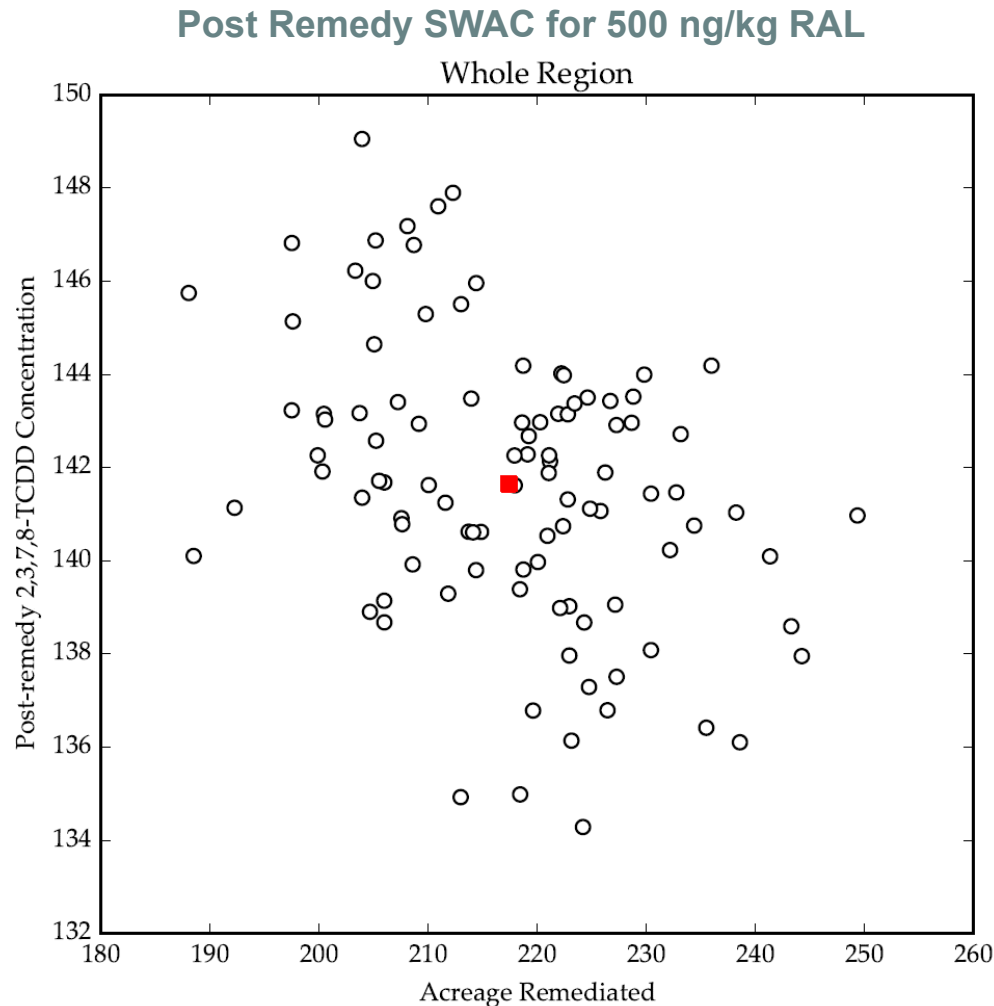
Fraction of Grid Cell Area Remediated =  $(0.33 + 0.67 + 0.67) / 3 = 56\%$

Model Grid Cell Conc Reduction =  $(1.37 - 0.23) / 1.37 = 83\%$

# Averaging of Simulations to 80-ft Decision Units



# Approach to Craft Remedial Options for FS



- Post-remedy SWAC vs acreage estimate is in the center of the simulation results

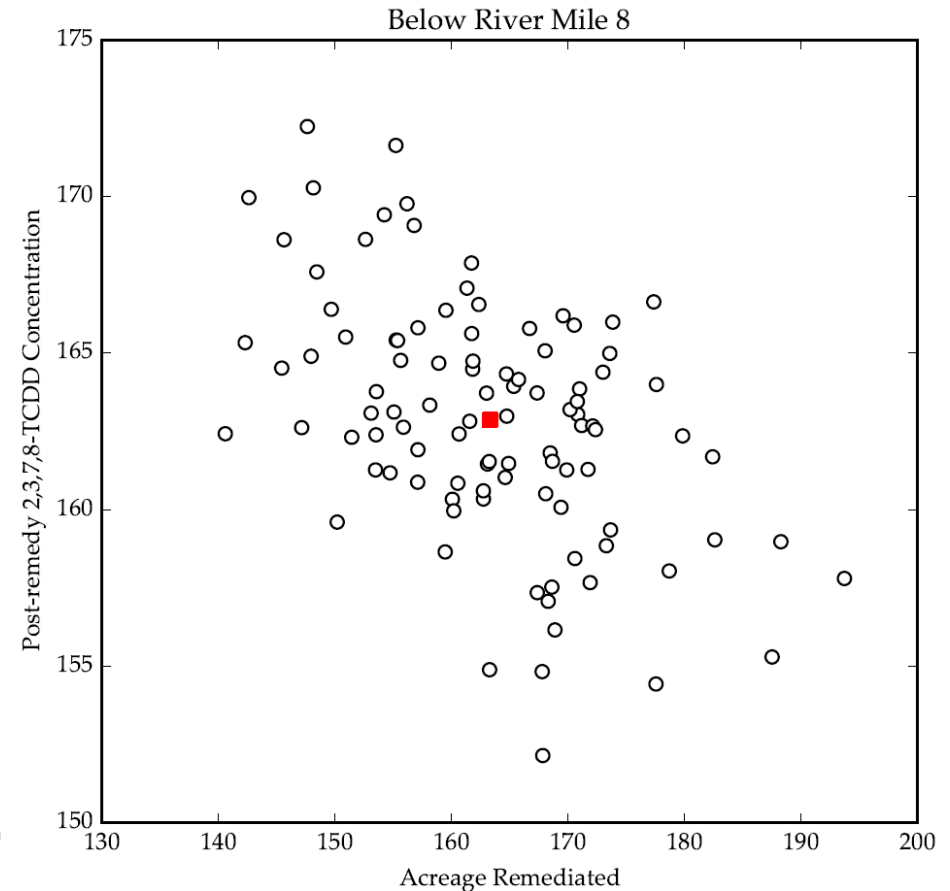
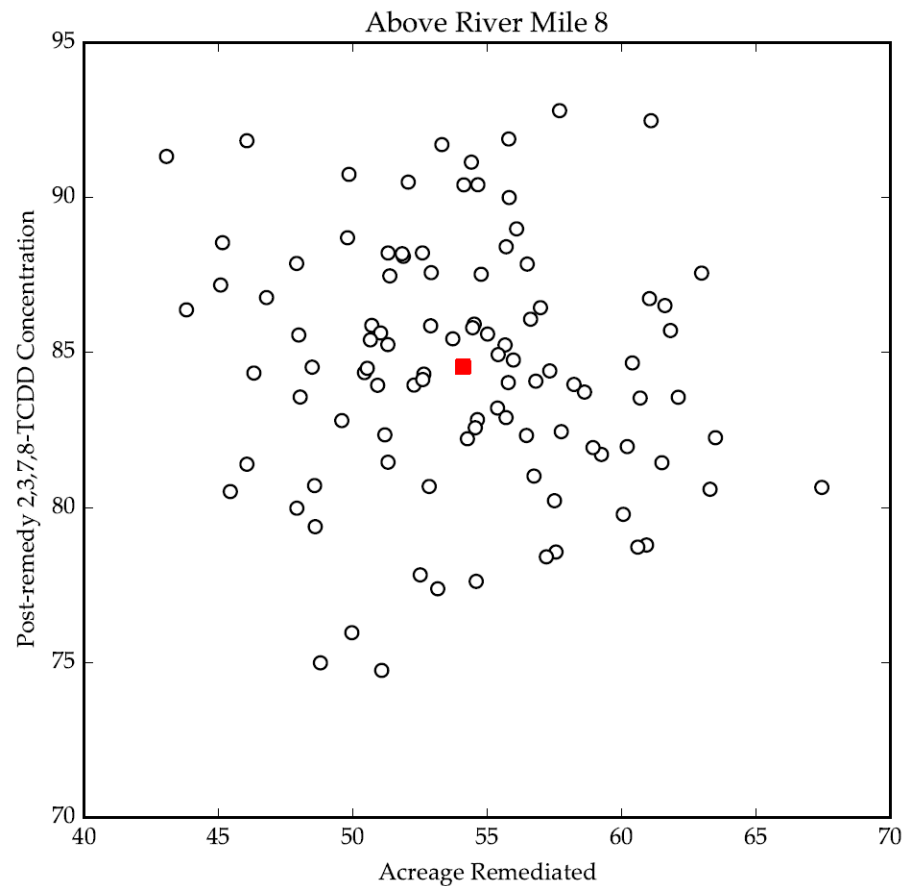
 Proposed Approach



# Approach to Craft Remedial Options for FS

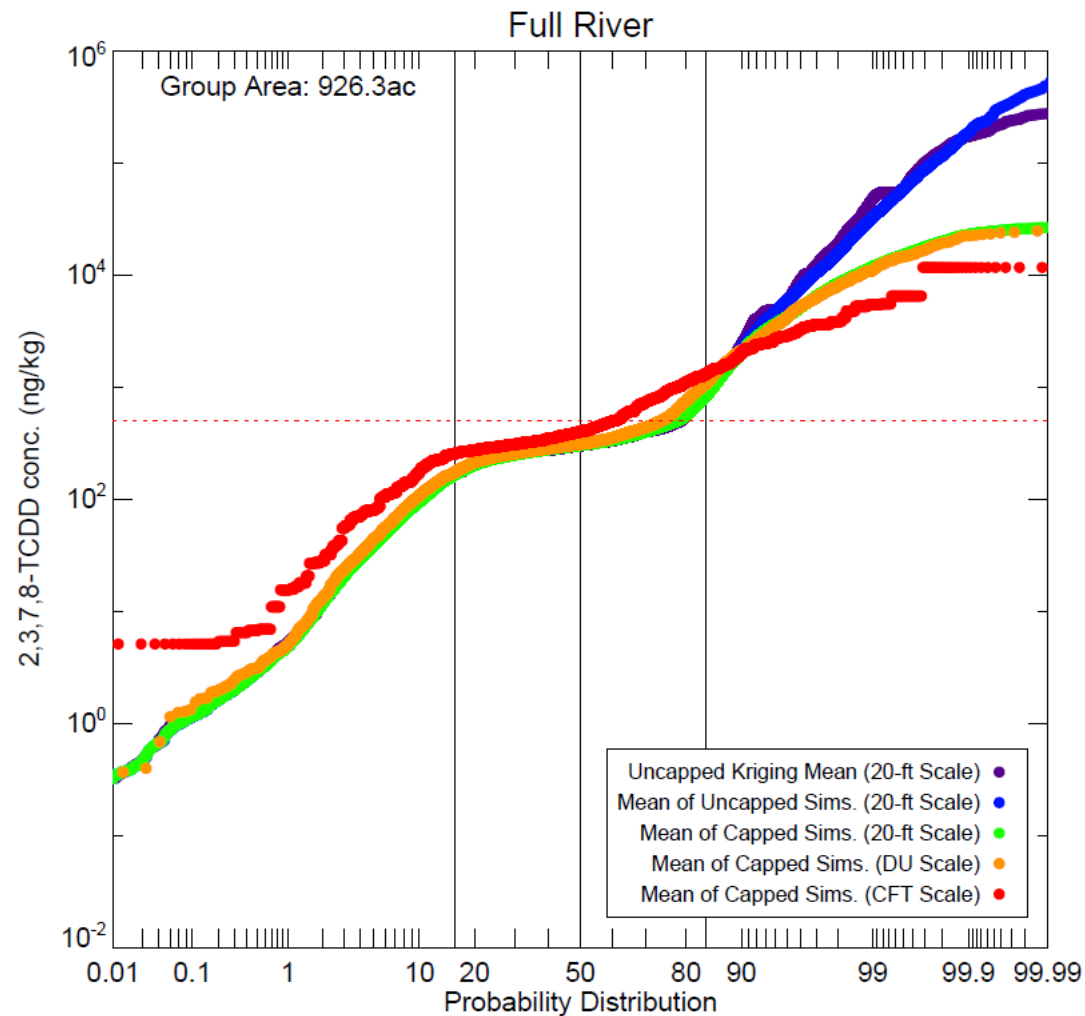
> RM 8 Post Remedy SWAC for 500 ng/kg RAL

< RM 8 Post Remedy SWAC for 500 ng/kg RAL



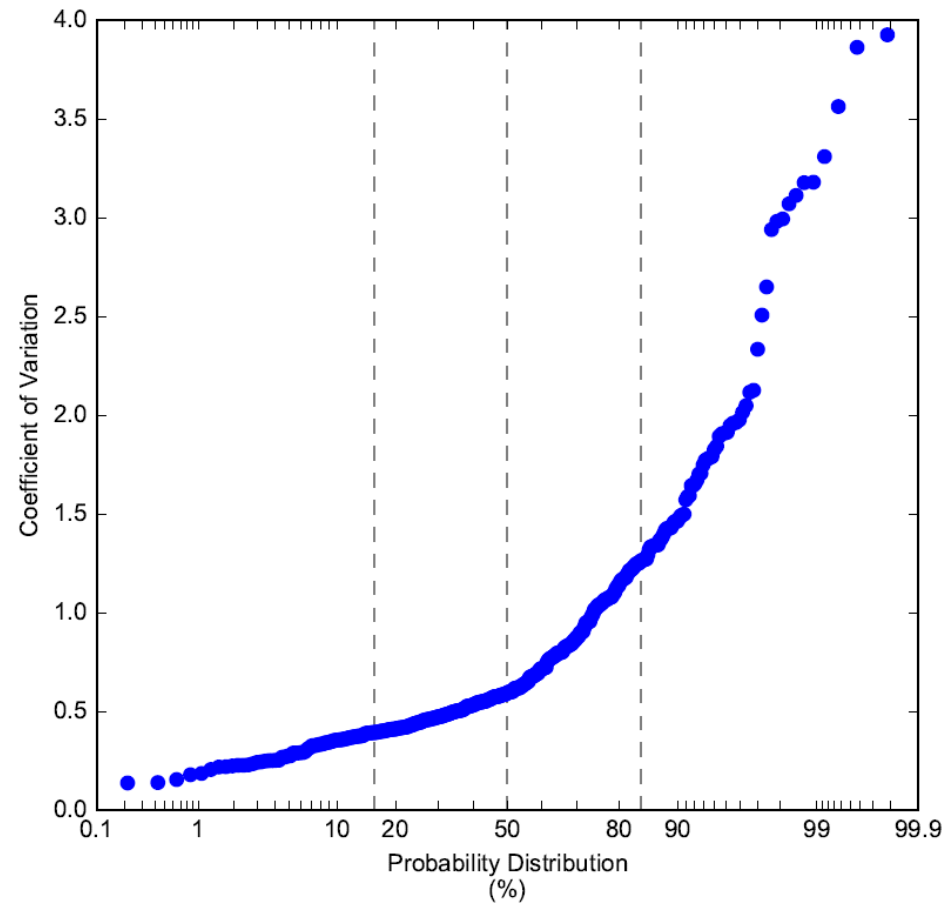
■ Proposed Approach

# Characteristics of Proposed Model IC

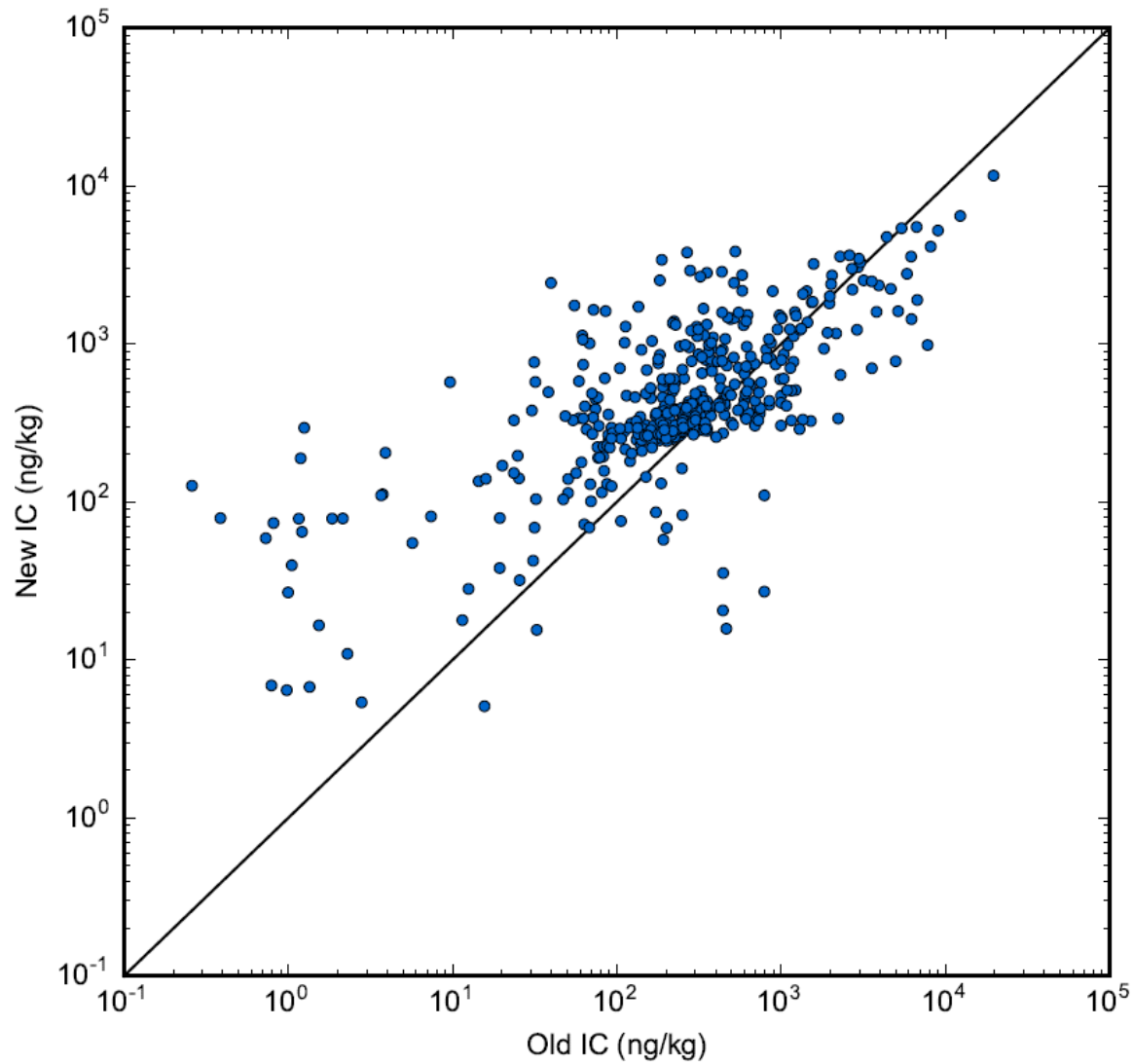


# Concentration variability across simulations on CFT grid scale

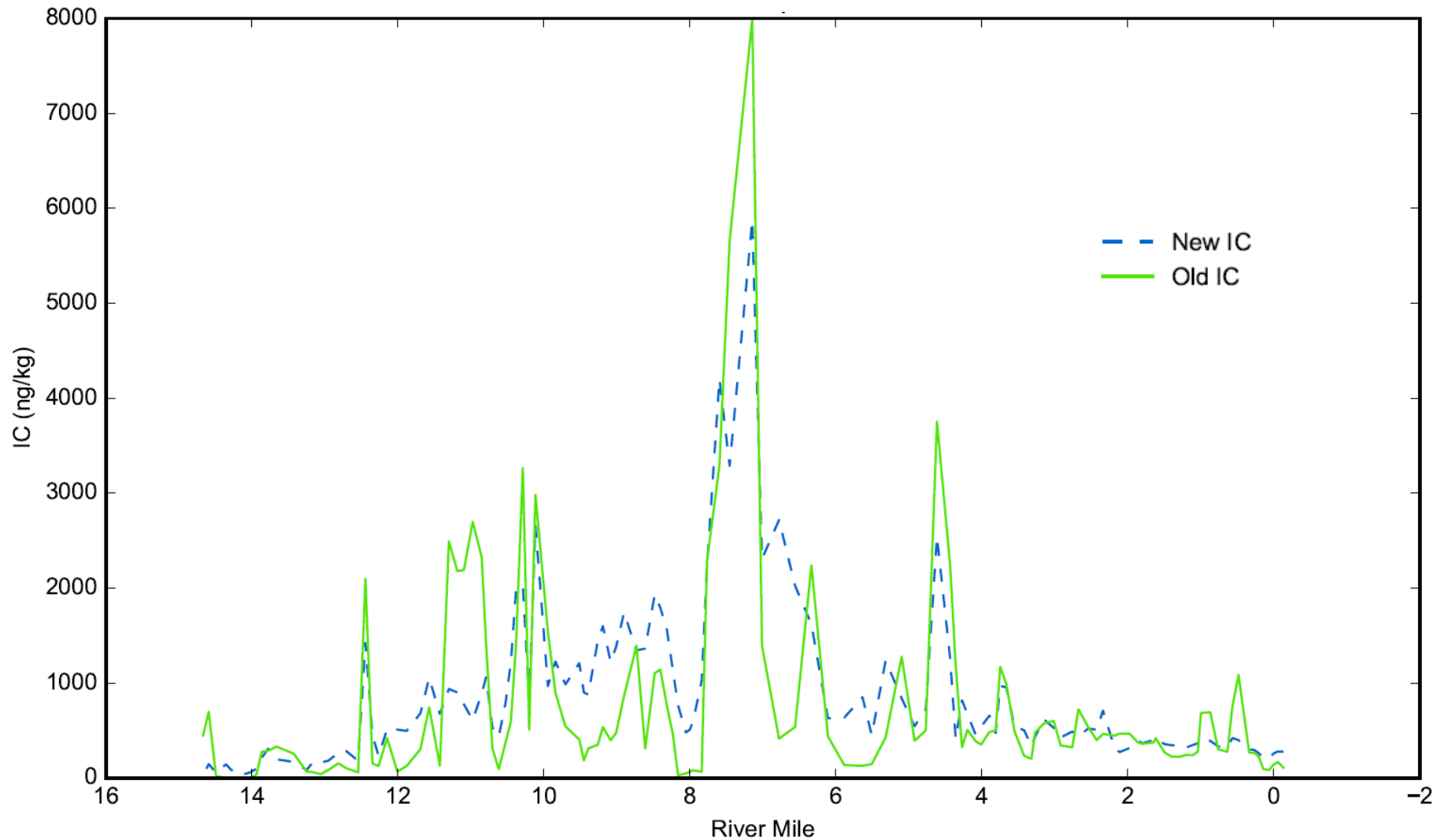
Coefficient of Variation from 100 Conditional Simulations for Each CFT Model Grid Cell



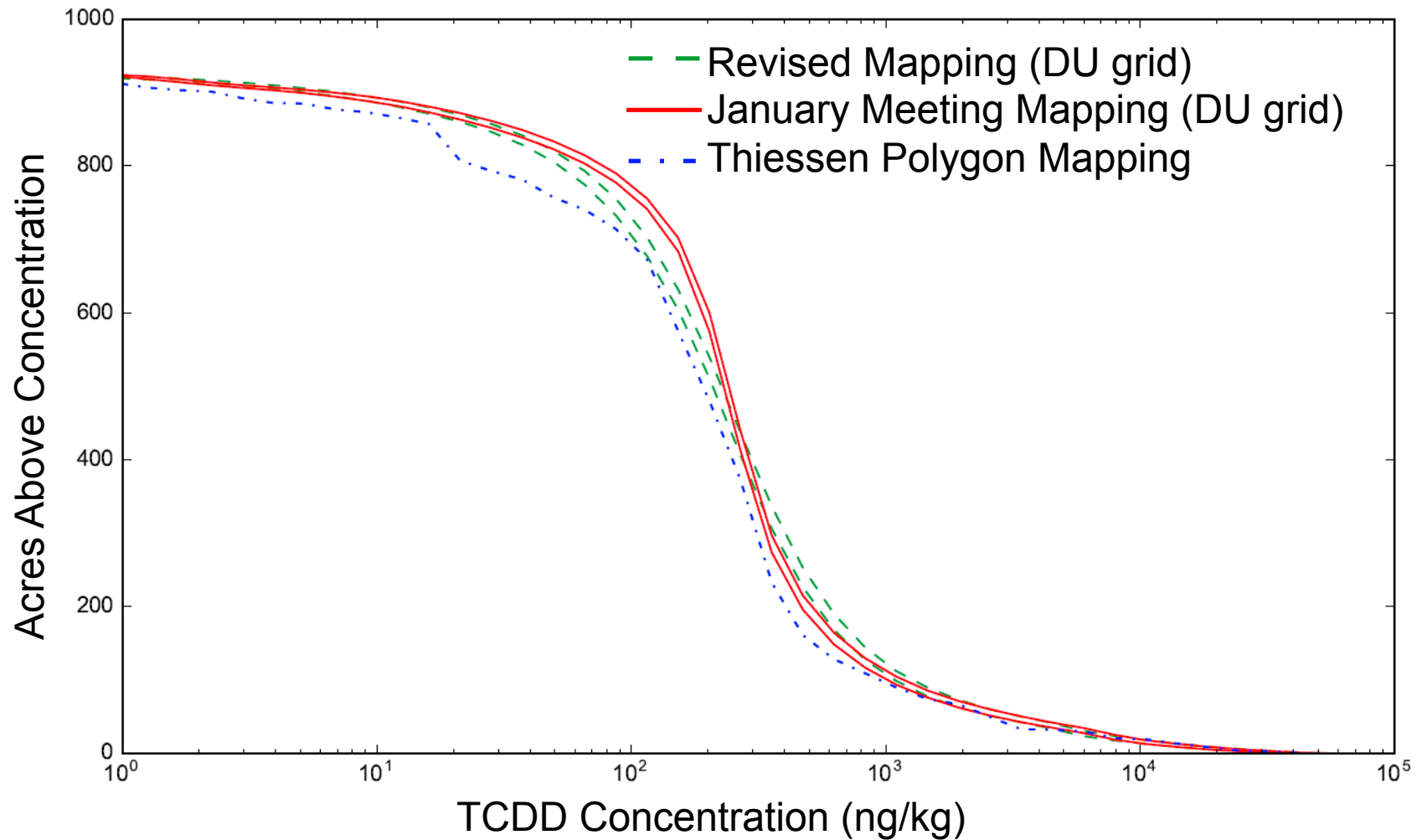
# Comparison of New and Old CFT Model IC



# Comparison of New and Old CFT Model IC

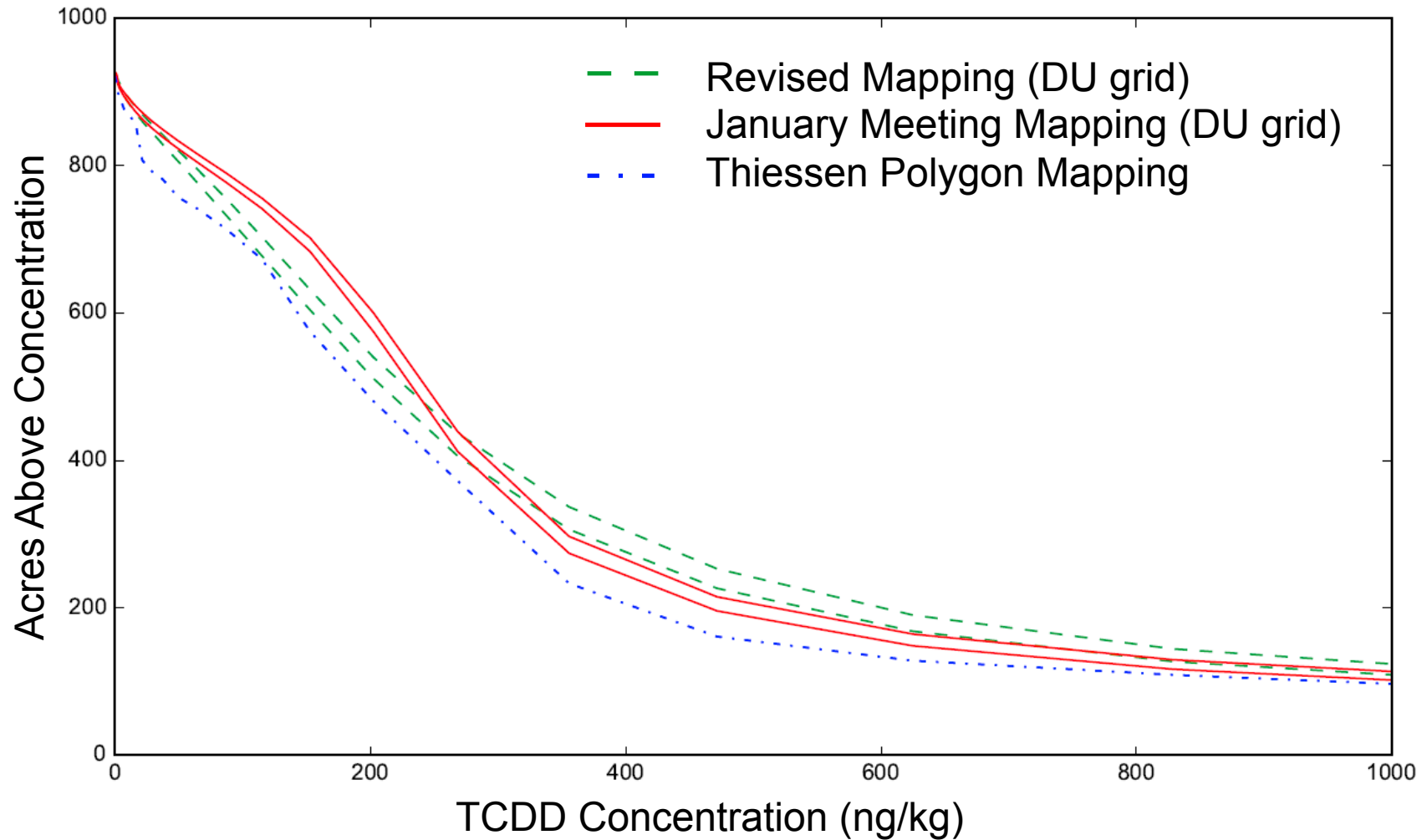


# Comparison of Mapping Results



*The lines represent  $\pm 2$  SD Envelope*

# Comparison of Mapping Results (Zoomed in)



*The lines represent +/- 2 SD Envelope*

# Summary: New Mapping vs Thiessen Mapping

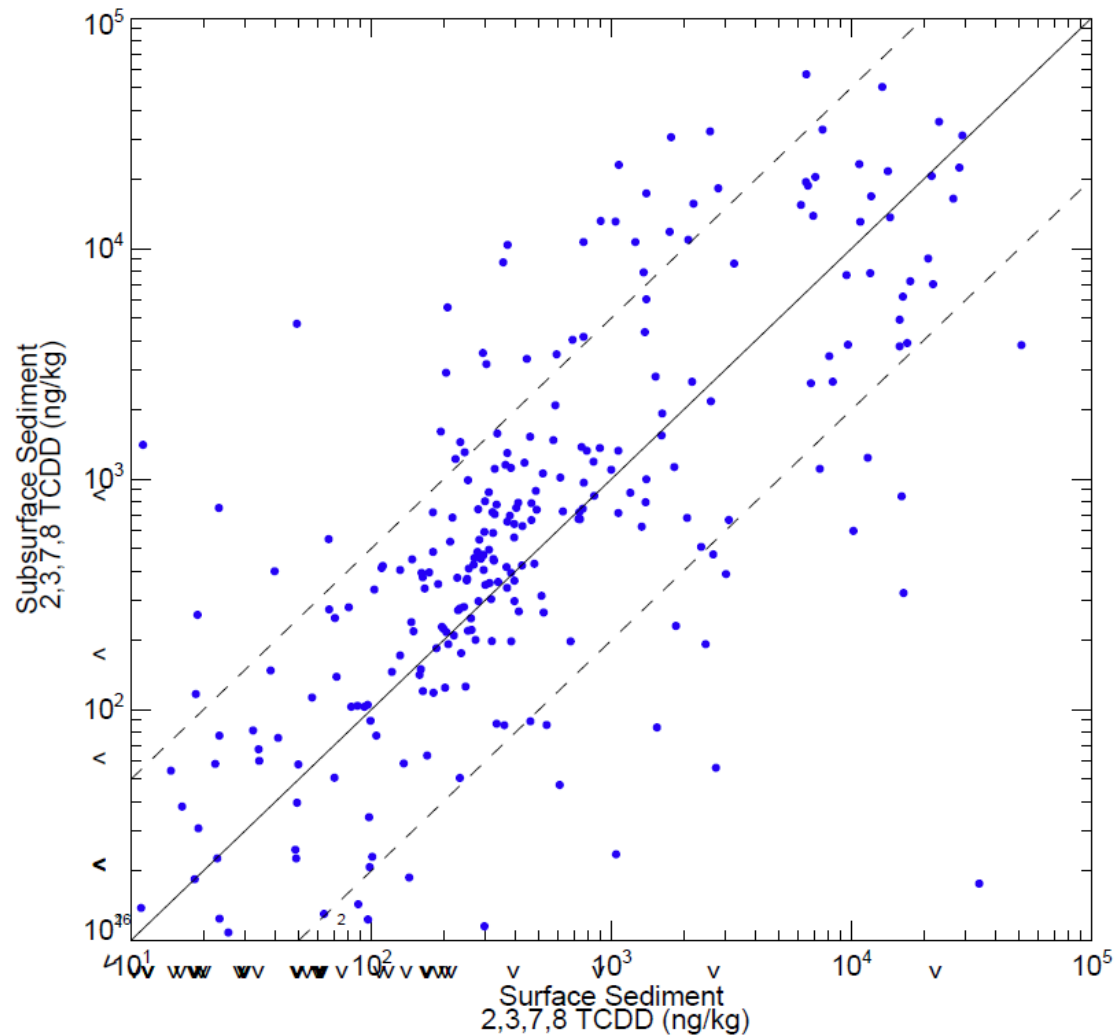
- Benefits of the revised mapping approach
  - Incorporates uncertainty into the mapping and the crafting of remedial alternatives for the FS
  - Provides an improved representation of concentration variability within the river
  - Yields a model IC that is smoother than the prior Thiessen-based IC



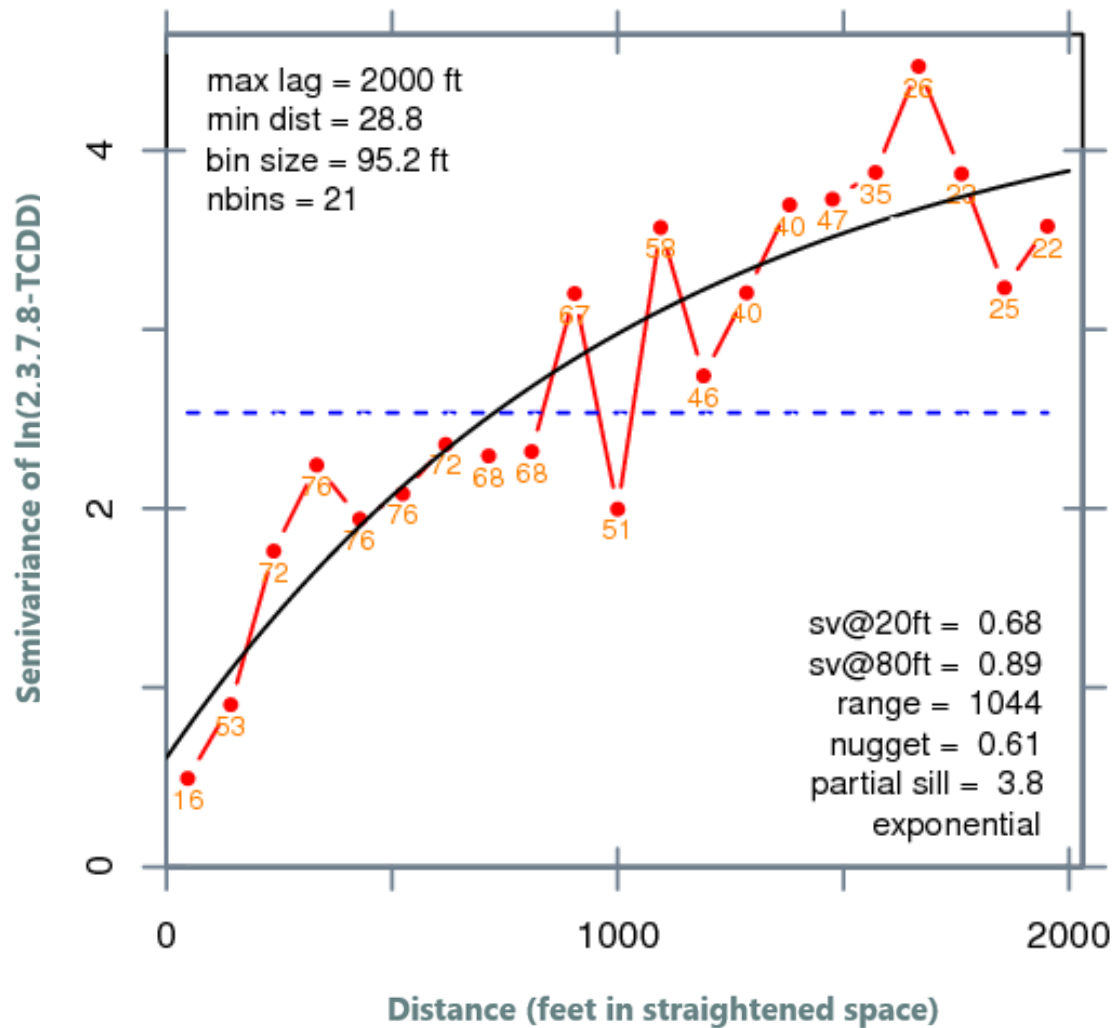
# Mapping Subsurface Concentration

- Evaluated the following
  - correlation between surface and subsurface concentration
  - variograms in the subsurface
  - suitability of channel bathymetric groupings for subsurface
- Propose Thiessen polygons as a practical option
  - FS evaluations will be less sensitive to subsurface mapping
    - Only influences predictions in erosional model cells
    - Not used in delineation of target areas
  - Thiessen mapping would likely avoid complications

# Subsurface vs. Surface Concentration Data



# Subsurface RM 10.9 Variogram



# Proposed Next Steps

- “Declare victory” – CPG memorialize approach; EPA rapidly approve
- Transition from mapping to revising the model based on EPA RI Report comments
- Extend mapping approach to other COPCs
- Apply refined maps to generate model initial conditions
- Revise model calibration as necessary based on refined mapping and RI comment-motivated changes to the models